MINISIS USERS' GROUP
MEETING 1989

IDRC
CRDI
CIID

CANADA

The International Development Research Centre is a public corporation created by the Parliament of Canada in 1970 to support research designed to adapt science and technology to the needs of developing countries. The Centre's activity is concentrated in six sectors: agriculture, food and nutrition sciences; health sciences; information sciences; social sciences; earth and engineering sciences; and communications. IDRC is financed solely by the Parliament of Canada; its policies, however, are set by an international Board of Governors. The Centre's headquarters are in Ottawa, Canada. Regional offices are located in Africa, Asia, Latin America, and the Middle East.

Le Centre de recherches pour le développement international, société publique créée en 1970 par une loi du Parlement canadien, a pour mission d'appuyer des recherches visant à adapter la science et la technologie aux besoins des pays en développement; il concentre son activité dans six secteurs : agriculture, alimentation et nutrition; information; santé; sciences sociales; sciences de la terre et du génie et communications. Le CRDI est financé entièrement par le Parlement canadien, mais c'est un Conseil des gouverneurs international qui en détermine l'orientation et les politiques. Établi à Ottawa (Canada), il a des bureaux régionaux en Afrique, en Asie, en Amérique latine et au Moyen-Orient.

El Centro Internacional de Investigaciones para el Desarrollo es una corporación pública creada en 1970 por el Parlamento de Canadá con el objeto de apoyar la investigación destinada a adaptar la ciencia y la tecnología a las necesidades de los países en desarrollo. Su actividad se concentra en seis sectores: ciencias agrícolas, alimentos y nutrición; ciencias de la salud; ciencias de la información; ciencias sociales; ciencias de la tierra e ingeniería; y comunicaciones. El Centro es financiado exclusivamente por el Parlamento de Canadá; sin embargo, sus políticas son trazadas por un Consejo de Gobernadores de carácter internacional. La sede del Centro está en Ottawa, Canadá, y sus oficinas regionales en América Latina, Africa, Asia y el Medio Oriente.

This series includes meeting documents, internal reports, and preliminary technical documents that may later form the basis of a formal publication. A Manuscript Report is given a small distribution to a highly specialized audience.

La présente série est réservée aux documents issus de colloques, aux rapports internes et aux documents techniques susceptibles d'être publiés plus tard dans une série de publications plus soignées. D'un tirage restreint, le rapport manuscrit est destiné à un public très spécialisé.

Esta serie incluye ponencias de reuniones, informes internos y documentos técnicos que pueden posteriormente conformar la base de una publicación formal. El informe recibe distribución limitada entre una audiencia altamente especializada.

IDRC-MR281e
March 1991

PROCEEDINGS
COMPTE RENDU

91-04061/

第11届国际MINISIS用户年会
11TH Annual Meeting
MINISIS Users' Group
11è Réunion annuelle
Groupe des Utilisateurs de MINISIS

MUG
89
BEIJING

机械电子工业部　机械科技情报研究所
Scientech Information Centre
Ministry of Machinery & Electronics Industry
Centre d'information scientifique et technique
Ministère de l'Industrie mécanique et électronique

中国，北京
Beijing, China
Beijing, Chine

1989年10月9-13日
October 9-13, 1989
9 au 13 octobre 1989

Material contained in the report is reproduced in the original language of presentation, English, Chinese, or French. The material is produced as submitted and has not been subjected to peer review or editing by IDRC Communications Division staff. Unless otherwise stated, copyright for material in this report is held by the authors. Mention of proprietary names does not constitute endorsement of the product and is given only for information.

Table of Contents

page #

1. MINISIS Version H and Beyond
Terry Gavin
Associate Director
WandaJane Phillips
Computer Documentation Specialist
International Development Research Centre 1

2. MINISIS APPLICATIONS IN STIC
Ms. Li, Linjia
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries 7

3. Application Development in Version H
Richard Lee 13
IDRC

4. APPLICATION AND POPULARIZATION OF MINISIS IN CHINA
By Jiang Xiangdong 21
(MINISIS Resource Centre in China)

5. MINISIS at Services documentaires multimedia (SDM)
Alain Boucher 29
Services documentaires multimedia (SDM)

6. MINISIS AT PID
D S R MURTY & CHANDER SHEKHAR 43
Publication & Information Directorate

7. Installation of Large Databases by Means of MINISIS
Wu Bing and Li Yanxia
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries 49

8. Establishment of APT Database on MINISIS
Yu Xinsheng 55
Louyang Petro-Chemical Engineering Corporation

9. FISH GERMPLASM RESOURCES DATABASE APPLICATION SYSTEM (FGRAS)
Ni Rufeng, Huang Nenggeng
Freshwater Fisheries Research Centre,
Chinese Academy of Fisheries Sciences 59

10. Establishment of Japanese Document Retrieval System Based on MINISIS DBMS
Wu Kang
Zhou Jiang
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries 69

11. Setting up Large Scale Document Searching Data Base with the MINISIS DBMS
LIU, YUN
The Scientech Information Centre
The Ministry of Machinery & Electronic Industries 75

12. Go Further into the Retrieval Capability of Specialized Data Base
Wu Bing
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries 85

13. THE APPLICATION OF MINISIS IN UNION CATALOG OF PERIODICALS
Wang Yan-mei
Library of the Chinese Academy of Medical Sciences (CAMS)
and Peking Union Medical College (PUMC)
Translated into English by Chen Cheng-pei 91

14. DATABASE OF CHINESE MECHANICAL ENGINEERING ABSTRACTS (DB CMEA)
Yang Zhe-zheng, Dong Chun-yuan, and Pei Xiao-ming
The Scientech Information Centre 105

15. Building Domestic and Foreign Instrument Products Data Base through use of Minisis
Wu Xiaohong and Sun MengLan
Instrumentation General Technology and Economics Institute
Ministry of Machinery and Electronics Industry 125

16. Application of MINISIS to Family Planning Material Retrieval
Liu Zhao-hui
Yoa Xiang-yang
Sichuan Population Information Centre 131

17. The Establishment of Computerized Agricultural Databases in Chinese Language
CHUNPEI HE
Scientech Documentation and Information Center
Chinese Academy of Agricultural Sciences 137

18. MARC Applications in Chinese Bibliographic Databases
Ms. Wang Ming
China Population Information Centre 141

19. Acquisition, Cataloguing and Public Access Subsystems of STIC Integrated Library System
Li, Linja
Shi, Minghong
The Scientech Information Centre 151

20. MINISIS IN THE CYRILLIC AND GREEK ALPHABETS AT THE MULTILINGUAL BIBLIOSERVICE OF THE NATIONAL LIBRARY OF CANADA 163
Virginia Ballance

21. The Periodical Computer Management System based on Combination of Chinese and Western Languages (PCMS)
JIANG WEI FEZ
ZHAO HAS JUN
Airforce Political College Computer Centre
Jiang Wan Wu Jiao Chang 167
Shanghai China

22. Warnier Methodology Applied to the Integral Design of a Library System Using MINISIS
M.en C. Ivonne Guzmán
Lic. Ofelia Solis
Lic. Ricardo González
Universidad Autónoma Metropolitana
Mtra Luz Marina Quiroga
Mtro Roberto Garduño 173
Universidad Nacional Autónoma de Mexico

23. MINISIS : An Experience in Health Sciences Library CAHSL Subsystem
TANG ZHAOBI
YUAN ZUPING
(CAHSL – Cataloguing and Acquisition Subsystem for the Health Sciences Library of Chinese Academy of Medical Sciences) 197

24. INTEGRATED SYSTEM OF PAO SUI-LOONG LIBRARY
ZHEN QIAO-YING
YAN CHUN 207
Pao Sui-loong Library of Shanghai Jiao Tong University

25. NATIONAL UNIVERSITY OF SINGAPORE LIBRARY OUTLINES OF LIBRARY MANAGEMENT SYSTEM 215
ONG GIN HONG

26. MINISIS VERSION H APPLICATIONS
BOB McKERCHER 219
IDRC

27. IDRIS - an Inter-Agency Development Research Information System
Mary Campbell
IDRIS Project Coordinator 225
International Development Research Centre

28. IMPLEMENT SYSTEM OF ULTRASONIC DIAGNOSIS MANAGEMENT ON THE MINISIS DBMS
Li Shilang
Yao HAO
Wang Zhe
Computer Room Library
China-Japan Friendship hospital, Beijing,China 235

29. SELECTIVE DISSEMINATION OF INFORMATION USING MINISIS AT THE NATIONAL INFORMATION CENTRE, S.N.D.T. WOMEN'S UNIVERSITY, BOMBAY
Ms. Harsha Parekh (Deputy Librarian - Co-ordinator) 249

30. The Chinese Data Entry with BATCHIN
Zhou Jiyu
Shanghai Population Information Centre 255

31. DEVELOPMENT AND MANAGEMENT OF AN AGRICULTURAL BIBLIOGRAPHIC INFORMATION SYSTEM
Franz Leemreize
Wageningen Agricultural University
The Netherlands 263

32. INCREASE INPUT SPEED OF CHINESE CHARACTER BY USE OF BATCHIN
Wu XiaoHong
Instrumentation General Technology and Economics Institute
Ministry of Machinery and Electronics Industry 271

33. Minisis in King Faisal Center
By Ibrahim Salem 285

34. Data Transferance Between CDS/ISIS and MINISIS
Presented by Xi Zhenyan
Written by Kin Zhenyan
Technical Services Office
Shanghai Population Information Centre 291

35. An Experience in SPEEDWARE-MINISIS
Mr. Lic. Ricardo Gonzalez Moreno
Cooridinacion de Servicios de Computo
UNIVERSIDAD AUTONOMA
METROPOLITANA–XOCHIMICO
MEXICO D.F. 313

36. ADVANTAGES OF USING MINISIS IN THE DEVELOPMENT
OF THE INFORMATION DATA BANKS FOR THE
PROJECT:
CULTURAL UNIVERSE OF THE STATE OF MEXICO
Josè Ricardo Montenegro G.
Unity of Informatics Services
I.M.C. 317

37. Multilingual Thesaurus Using MINISIS
Sun Minglei
Zhou Jiyu
Shanghai Population Information Center 325

38. The Implementation of the SNOMED System using the MINISIS
DBMS in Chinese-English
Wang Zhe
Ding Zhöian
Library
China-Japan Friendship Hospital
Computer Centre, ScienTech Information Centre
The Ministry of Machinery-Electronics Industry 343

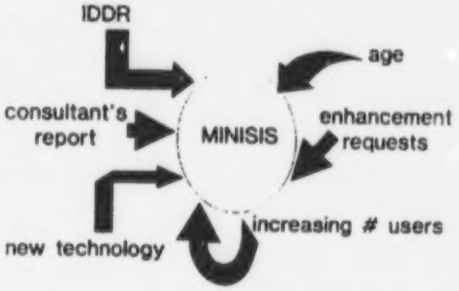
39. COMPUTERISATION AT THE NATIONAL LIBRARY OF
MALAYSIA:
FUNCTIONAL SPECIFICATIONS FOR AN
INTEGRATED LIBRARY SYSTEM
Lim Chiew Aun
National Library of Malaysia 353

40. MINISIS Version H – Documentation
WandaJane Phillips
IDRC 363

MINISIS Version H and Beyond

Terry Gavin
Associate Director
WandaJane Phillips
Computer Documentation Specialist
International Development Research Centre

BACKGROUND



IDDR
age
consultant's report
MINISIS
enhancement requests
new technology
increasing # users

RESPONSE

MINISIS
Version H.00

OBJECTIVES

- responsive to requirements of MINISIS user community
- responsive to changes in technology
- produce version of MINISIS to operate on variety of hardware

MINISIS ACTIVITIES

3 separate but related activities

- technical development
- dissemination and support
- documentation and training tools

TECHNICAL DEVELOPMENTS
General

- improve user interface
- enhance application development tools
- migrate software to C

TECHNICAL DEVELOPMENTS
RESULTS

- increased flexibility and additional capabilities for applications development
- possibility of decentralizing technical support and development
- operation of software on other hardware

TECHNICAL DEVELOPMENTS
SPECIFIC

- improved user interface
    - menu
    - command mode
- specific turnkey applications
- better support for full text
- enhanced numeric processing
- applications development language

TURNKEY APPLICATIONS

- Generalized Application
    - replacement for MINISIS Version G.01
- IDRIS Application
    - model for project information systems
- Integrated Library Systems
    - will include acquisitions, circulation, serials control and enhanced MARC support

HARDWARE
Minicomputer Environment

- any model of HP 3000 - from Micro 3000 to Series 960
- configuration depends upon applications (MINISIS and others) and # of users to be supported

HARDWARE
Microcomputer Environment

- IBM PC/AT or larger including 100% compatibles
- minimum 640 Kbytes RAM with LIM specification support for expanded memory
- monochrome display
- minimum 5 Mbytes hard disk space PLUS free space to store and manipulate data files - ratio 2:1
- 1 or more 3.5 or 5.25 floppy diskette drives
- MS or PC DOS Version 3.3 or higher
- printer with alternate character set capability, if it is a requirement of the application

HARDWARE
Other Hardware Environment

- will be developed with IDRC assistance
- 80-85% of code will be common to all hardware
- 3rd party to develop specific machine dependent routines
- software to be provided under same terms as provided by IDRC
- IDRC to maintain common code
- 3rd party to resolve hardware-specific problems

MINISIS RESOURCE CENTRES

- promote MINISIS
- introductory and advanced training
- support and trouble shooting
- development of application tools
- production and maintenance of local language documentation
- test site for new versions

- program to establish network already underway
- currently
  - China
  - North Africa/Middle East
  - India
  - West Africa

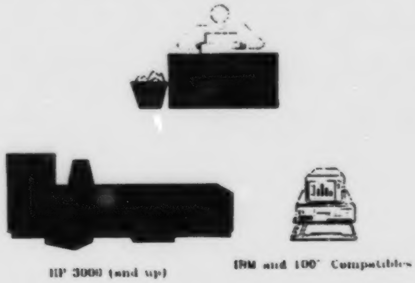
- 1989/90
  - Mexico
  - China - phase II
- 1990/91
  - Southeast Asia
  - NA/ME - phase II
- 1991/92
  - East Africa
  - India, West Africa - phase II

DOCUMENTATION and TRAINING

- closely related to technical developments and dissemination & support
- recognize current documentation dense, technical and unfriendly
- need to develop self-training tools

MINISIS DOCUMENTATION

The Expanding MINISIS Constituency



Documentation for Version H
Process of Development

ANALYSIS	DESIGN	DEVELOP/ VERIFY
• USERS • PRODUCT	• LOGICAL • PHYSICAL	• COMPLETE • ACCURATE • USABLE

END USER/TERMINAL OPERATOR
Version II Task List

- select database(s)
- search database(s)
- enter, select, and modify record(s)
- create and edit specifications for: PRINT, INDEX, COMPUTE
  list and manage database specifications
- format and print records
- sort records
- COMPUTE
- RENUM

DATABASE AND SYSTEM MANAGER
Version H Task List

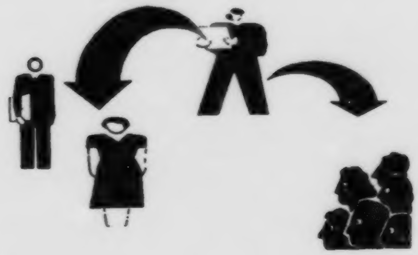
- all end user activities
- import and export database definitions
- create and edit database definitions
- inverted file management
- database maintenance
- user and security profile definitions
- application customization

- statistics
- list specifications of user profiles
- diagnostics
- recovery

PROGRAMMER
Version H Task/Information List

- system information
- machine dependent routines
- implementing alternate character sets
- interfacing MINISIS with user defined data structures
- 3GL programming using the High Level Intrinsics

USER QUESTIONNAIRE



USER QUESTIONNAIRE
Assumptions

QUESTIONNAIRE RESULTS

- information is thorough ('fairly complete')

- information is fragment
- information is difficult to find
- concepts and language difficult

TRAINING MATERIALS

Concepts

- recent studies show a method called MENTAL MODELS or CONCEPTUAL MODELS assist users in learning

Users with accurate and complete mental models of a system are better able to learn new commands and

Exercises

- concepts backed by concrete exercises become 'second nature' for users.

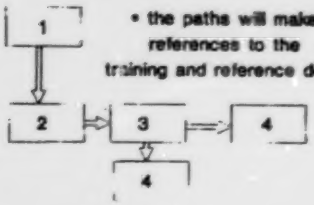
The combination of mental model, or representation, with physical activity seems to offer the best system of learning...

REFERENCE MATERIAL

- the information currently available needs to be restructured for easy access
- the help screens need to be redesigned to reflect similar changes in the reference manual
- the help screens, the training materials, and the users' guide need to direct users to the complete information in the reference guide

USERS' GUIDE

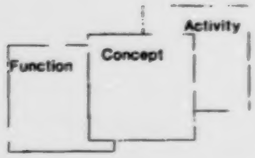
- will present common operations



- the paths will make references to the training and reference docs

DEVELOPMENT

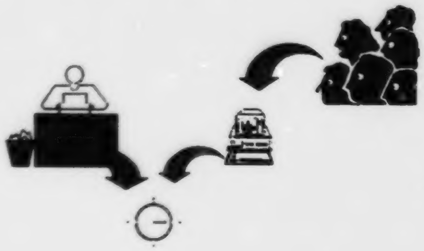
Modular Design



- task oriented
- problem solving
- modular
- conceptual

the functionality of Version H is logically separated and each logical unit written separately

USABILITY STUDIES



CONCLUSIONS

- natural next step in evolution of MINISIS
- functionally similar to Version G.01
- restructuring internals and user interface
- foundation for new developments and applications in subsequent versions

MINISIS APPLICATIONS IN STIC

Ms. Li, Linjia
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries

This article aims to give our colleagues a general review on the development of MINISIS applications in STIC, so as to exchange lessons and experiences, and seek advice from them.

Brief Introduction of STIC
— the General application environment of MINISIS

STIC, is an integrated part of the information complex STIP — the Scientech Information & Publications, which is a scientific and technological information and publishing department under the Ministry of Machinery & Electronics Industries. There are two branches in STIP, the Scientech Information Centre (STIC) which was established in 1958, and China Machine Press (CMP) which was established in 1952, and the two organizations merged into STIP in 1970. The name of "STIC" is still in use, but it generally refers to the STIP divisions involved in the area of information. The Central Library for Mechanical & Electrical Engineering is a group of such divisions.

STIC's computer system is composed of two HP computers, the HP3000 / III and HP3000 / 68, as well as some micro-computers. The HP3000 III was installed in 1981, with 1 MB main memory storage capacity and 480 MB of external storage capacity. While the HP3000 / 68 was installed in 1985, with a main memory storage capacity of 4 MB, and an external storage capacity of 5112MB.

MINISIS Applications

MINISIS was first installed in our computer system in 1981. That was the first time that MINISIS staged in China. Since then, its application projects in STIC have grown considerably. In general, the applications can be categorized into three types: database-oriented, library-oriented, and thesaurus-oriented.

In the first period, from 1981-1984, owing to the lack of Chinese character processing function, MINISIS was mainly used to establish documentary databases in Western languages, such as ISMEC (Information Service in Mechanical Engineering) and COMPENDEX (the machine-readable version of Engineering Index). By the end of 1983, the library applications in processing materials in Western languages were started.

For the period from 1985 up to now, since the technical environment for Chinese character processing was improved, and our computer storage capability extended, as well as our staffs became more familiar with MINISIS functions, MINISIS applications therefore were developed in a much wider scale.

Databases either in Western languages or in Chinese language have been established during this period. Those databases are:

INSPEC	the machine-readable file corresponds to the printed Physics Abstracts, Electrical and Electronics Abstracts, and Computer and

	Control Abstracts; which is issued by The Institution of Electrical Engineers, London, England.
COMPENDEX	the machine-readable version of the Engineering Index; which is issued by Engineering Index Inc., New York, U.S.A.
METADEX	the METALS ABSTRACTS ALLOYS INDEX, which is produced by American Society for Metals, Ohio, U.S.A.
CMEA	the machine-readable version of Chinese Mechanical Engineering Abstracts; which is produced by STIC.
JICST	the machine-readable version of JICST printed abstracts journal, which is provided by Japan Information Centre of Science and Technology.
EMA	the Engineering Material Abstracts, which is provided by American Society for Metals, Ohio, U.S.A.

The total number of records available in these MINISIS databases is about 1,830,000. Most of these databases can be considered as the large-scaled databases with MINISIS, and have already been in the stage of providing tentative services; some of them can be even accessed through remote terminals outside of Beijing city.

The applications developed in the library, during this period, are as follows:

Acquisition and Cataloguing System. It was developed to process monographs and documents in Western languages, and now its data is to be transferred into the new integrated system.

Online Foreign Serials Retrieval System. It was developed to collect all the information of serials in Western languages (foreign journals) held in the library, and provide search functions through some inverted keys, such as titles, abbreviations of title, journal codes, ISSN, etc.

The STIC Integrated Library System Planning. It was worked out on the basis of the library system analysis, and the feasibility studies; and provides a guideline for all the library application projects.

Acquisition, Cataloguing and Public Access Subsystems of the Library's Integrated System. They were set up in processing monographs and documents in both Chinese and Western languages.

In order to meet all the requirements of the application projects, our staff have not only practised almost all the function modules provided by MINISIS, but also developed a group of software:

Japanese Character Set. That was developed on the basis of the HP Chinese Character Set in supporting of establishing JICST, the Japanese documentary database.

Account,a MINISIS handler. It enables databases to render bills for users' access online.

Batchin Programs. They were developed to load those machine-readable information sources not in ISO 2709 format into MINISIS databases.

Data Transfer Programs. They allow the two-way data resource transfer between MINISIS database and either CD-ISIS, or DBASE Ill database.

Typesetting Program. It was developed to meet the requirements of computer typesetting for CMEA printed version; and permits adding the typesetting commands into MINISIS output files automatically.

Besides the projects we have completed there are application projects being undertaken and going to be developed. They are much more of library-oriented thesaurus-oriented applications, such as:

The subsystems of the library circulation and serials control, are in the stage of database structure and program design.

Thesaurus of METADEX is in the stage of batching in the data; while Thesaurus of EMA, and Thesaurus of Chinese Mechanical Engineering are in the stage of database design.

The Multilingual Terminology Databank of Mechanical & Electrical Engineering which is the Ministry of Machinery & Electronics Industries designated project has been decided to realize with MINISIS, and so is the National Terminology Databank, which is a project designated by the State Commission of Science and Technology.

A database of R & D results in machinery industry is already organized, and will be soon set up following the patten of CMEA. And another INSPEC database in its shortened version is going to take shape.

Another very important project our staffs are working hard on, is a **process for MINISIS Chinese character sorting.** We are trying to set up the sorting tables, first of all in Pinyin's alphabetical order, then in strokes' order, and later in orders of other Chinese character attributes. If we succeed, it is believed, all our peers with Chinese language applications will benefit.

Reviewing the past and looking into the future, we can see clearly that MINISIS has become an indispensable tool of STIC.

In closing, I'll take this opportunity to thank the staff from IDRC for their efforts in developing MINISIS, and for their warm-hearted help to us. Also I'll thank our colleagues, whoever helped us in the past. And our STIC staff are looking forward to exchanging experience with you, and will always be grateful for your kind advices.

MINISIS 在机电部机械科技情报所的应用

机电部机械科技情报所软件部
李林甲

这篇文章对于 MINISIS 在机械科技情报所的各方面应用进行了回顾，目的在于使同行对我们的工作情况有所了解，进而促进对于 MINISIS 应用方面经验教训的交流。

机械科技情报所梗概

——MINISIS 应用的环境

机械科技情报所，是机械电子部下属的一个科学技术情报和出版中心，它由两大部分组成——科技情报所和机械工业出版社。情报所建于 1958 年；而机械工业出版社建于 1952 年。情报所的概念通常是指那些在我单位从事情报研究和情报服务的部门。

情报所的使命是，根据部里重点任务的要求，收集、研究、分析、提供国内外机械科学技术的战略和战术情报。情报所也是机械工业全行业的科学技术信息和文献中心、以及咨询服务中心。它还负责组织和集中管理机械工业全行业的各情报信息网站和情报中心。另外机械中心图书馆是机械科技情报所对公众提供情报信息服务的窗口。

机械科技情报所拥有两套惠普小型计算机——HP3000／Ⅲ和 HP3000／68 型。还有一些微型计算机。1981 年情报所引进了 HP3000／Ⅲ计算机，拥有 1MB 的主存和 480MB 的外存容量。而 HP3000／68 机是 1985 年引进的，拥有内存 4MB，外存 5112MB。

MINISIS 应用项目

1981 年 MINISIS 首次来华就安装到我们的计算机系统上。它在机械科技情报所的应用项目不断开发，至今已有八年的历史。总的来说，MINISIS 的应用项目可大体分为三类：数据库建设、图书馆应用以及词表建设。

在 MINISIS 应用的第一个阶段里，从 1981 年至 1984 年，由于受汉字处理能力的限制，MINISIS 主要用于建立西文文献数据库，如 ISMEC (英国机械工

程文摘) 和 COMPENDEX (美国工程索引)等．到 1983 年底开始着手摘图书馆西文资料的处理．

从 1985 年至今, 由于汉字处理的技术环境改善了, 我们计算机存储容量扩展了, 也由于我们的工作人员对 MINISIS 更加熟悉了, MINISIS 的应用项目得到了蓬勃的发展．

在这一时期我们开展的数据库建设项目有:

INSPEC	与英国伦敦出版的《物理文摘》、《电气电子文摘》以及《计算机控制文摘》相对应的机读文档．
COMPENDEX	美国纽约的工程索引公司出版的机读版工程索引．
METADEX	美国金属学会出版的机读版《金属文摘和合金索引》．
CMEA	本所编辑加工的《中国机械工程文摘》机读版．
JICST	日本科技信息中心出版的文摘期刊机读版．
EMA	美国金属学会出版的工程材料文摘机读版．

这些 MINISIS 数据库的总装机容量已为 183 万篇记录．其中有些数据库可被认为是 MINISIS 的大型数据库, 并且部分已投人试运行和服务．有些数据库还连通了远程终端, 可供北京以外的用户检索．

这一时期开展的图书馆的应用项目有:

西文图书文献的采购编目系统　这一实验系统为新的联合系统的实施提供了良好的经验, 现已准备将其数据转人新的系统．

国外期刊联机检索系统　这一系统收集了馆藏国外期刊(西文)的各方面信息, 可通过诸多的倒排字段提供对馆藏期刊的各种检索, 如用标题、缩写刊名、正副刊名、ISSN 号等．

机械中心图书馆计算机联合系统总体规划　该规划建筑在对图书馆业务的全面分析和可行性研究的基础上．为实现图书馆的计算机应用制定了纲要．

采购、编目、检索子系统　这是本馆计算机应用联合系统中的一个率先实现的部分．它是一个综合处理中西文图书文献(一、二次文献)的应用系统．

为了满足各种应用项目的要求, 我们的工作人员, 不仅实践了几乎所有的 MINISIS 功能模块, 另外还开发了一些配合 MINISIS 使用的软件:

日本方块字字符集　它以 HP 计算机系统支持的汉字字符集为基础, 用于日本科技信息中心的文摘期刊数据库——JICST．

记帐程序　它是 MINISIS 的一个接口软件, 可允许用 MINISIS 建立的数据库联机向该库的用户提供帐单．

批处理程序　为了将非 ISO 2709 格式的信息资源装人 MINISIS 数据库, 我们的工作人员写了不少符合 MINISIS BATCHIN 模块功能要求的数据格式转换程序．

排版软件　该排版程序可将 MINISIS 中国机械工程文摘数据库的输出文件自动排序并加好下一工序计算机排版所需的指令, 以满足出版书本式《中国机械工程文摘》的需要．

除了上述我们已完成或接近完成的 MINISIS 应用项目以外，下面来谈一下我们正在进行和将要开发的 MINISIS 应用项目，它们大都属于图书馆应用的子系统或是词表建设系统．

图书馆流通管理和期刊管理子系统　这些系统正处在数据库结构设计及程序设计的过程中．

METADEX 主题词表　已完成模式设计，正在进行批处理装库，不久将为金属文摘数据库提供一条联机检索新途径．EMA 主题词表和《中国机械工程主题词表》正处在模式设计和数据格式转换的过程中．

多文种机电工程术语库　是机电部指定的词表建库项目，现已决定以 MINISIS 为基础来完成．国家序词库是国家科委指定的大项目，也已经确定由我所的工作人员以 MINISIS 为基础来完成．

科技信息成果数据库　是本所组织的情报管理应用项目．现已决定采用与 CMEA 相同的方式结合 MINISIS 来实现．

另外一个最重要的，我们的工作人员正在为之努力奋斗的项目，是 MINISIS 的汉字排序功能．我们准备首先建立按汉语拼音字顺方式排序的排序表，然后建立按汉字部首排序的排序表，最后建立汉字其它属性的排序表．如果我们能得到成功，这一成果将受益于广大有汉字应用项目的同行们．

回顾过去展望未来，我们可以清楚地看到，MINISIS 已经成为我们机械科技情报所事业中不可缺少的工具．

最后，我愿意利用这一机会向 IDRC 的朋友们表示感谢，没有他们的热情帮助我们不可能取得这些成绩．我还要向曾经帮助过我们的同行表示感谢．我们机械科技情报所的 MINISIS 工作人员，热切期望与大家交流经验，并欢迎同行们随时指点．

Application Development in Version H

Richard Lee
IDRC

What is a MINISIS Application Program?

- self-contained set of codes
- performs a specific application
- written in a high level programming language(s)
- uses MINISIS intrinsics
- examples:
    - MINISIS QUERY processor
    - CARDEX processor
    - CHECKIN processor

What is a MINISIS Exit?

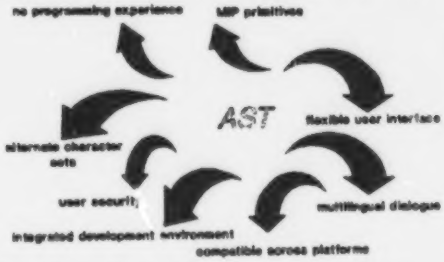
- a piece of executable code
- called by a MINISIS processor
- performs a user-defined function
- written in a high level programming language(s)
- optionally uses MINISIS intrinsics
- examples:
    - ROMAN8
    - TIMESAVE
    - MOD11CK

Considerations in the development of programs/exits in Version G

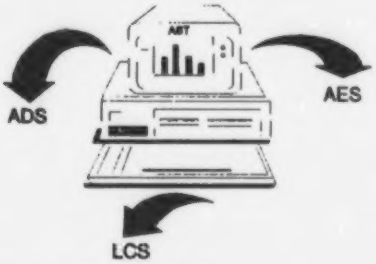
- knowledge of high level language(s)
- knowledge of MINISIS intrinsics
- maintain consistancy in user dialogue
- programmatically handle alternate character sets
- insufficient development facilities
- difficulties in code sharing
- limited user security

Application Specification
Toolbox

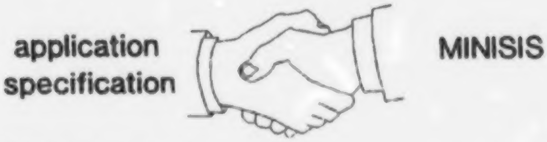
Application development in Version H


no programming experience
MIP primitives
AST
flexible user interface
alternate character sets
user security
multilingual dialogue
integrated development environment
compatible across platforms

Application Specification Toolbox


AST
ADS
AES
LCS

Subsystems of the AST

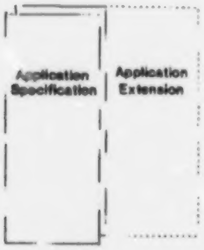

application specification
MINISIS

MIP Primitives

- no programming experience
- non-procedural command language
- SQL commands with extension
- code sharing

MIP primitives by category

- Data Manipulation Language (DML)
- Data Definition Language (DDL)
- Data Control Language (DCL)
- Data Processing Language (DPL)


Application Specification
Application Extension

Application Extension

- no programming experience
- extend and/or redefine Application Specification
- customize user dialogue
- multiple Application Extensions share an Application Specification
- prevent corruption of Application Specification
- compatible across platforms

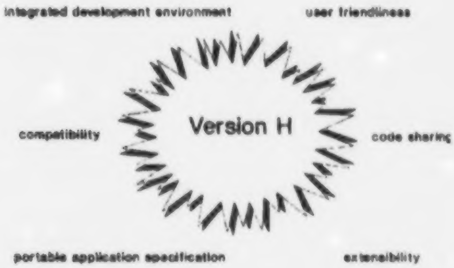
Developing Application Programs in Version H

- MINISIS manages installation of Application Program
- written in high level programming language(s)
- using MINISIS Version G as well as Version H intrinsics
- MINISIS Version H intrinsics provide users with better user interface routines
- user security is tied to the Application Program
- may not be transferable across platforms

Developing an Exit in Version H

- MINISIS manages installation of exit
- written in high level programming language(s)
- called by MINISIS internals, Application Specifications and Applications Programs
- optionally using MINISIS intrinsics
- may not be transferable across platforms

Conclusion


integrated development environment
user friendliness
compatibility
Version H
code sharing
portable application specification
extensibility

AST

Design Philosophy

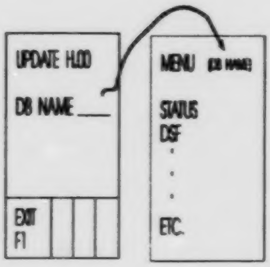
- no programming experience eg. Variables declaration, procedures declarations, etc.
- easy-to-define and flexible user interface eg. command, menu, window etc
- code sharing
- code with commands rather than language instructions
- support multilingual dialogue
- support customization of application

Layout

- to display and/or prompt for information
- consists of message(s) and input area(s)
- use WYSIWYG (What You See Is What You Get) approach to creating layouts
- define attributes to prompts (inputs) eg. editing rule(s), holding information, mapping information etc.
- define syntax of input
- define relationship among prompts
- define user security

Examples:

```
UPDATE VERSION H.00
ENTER DB NAME OR EXIT – XXXX
Q{STATUS|DSF|...etc.}
```

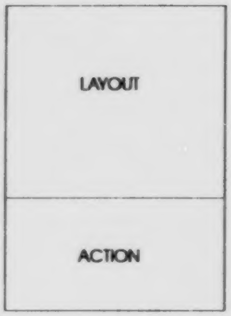

UPDATE H.00
DB NAME ___
EXIT
F1
MENU (DB NAME)
STATUS
DSF
.
.
.
ETC.

Action

- consists of a set of commands
- tied to a layout
- a command is either an MIP primitive or a link to other layout(s)

For example:

```
CALL “LAYOUT”,
SELECT * FROM DB1
WHERE ISN = 1,
UPDATE * INTO DB1
```

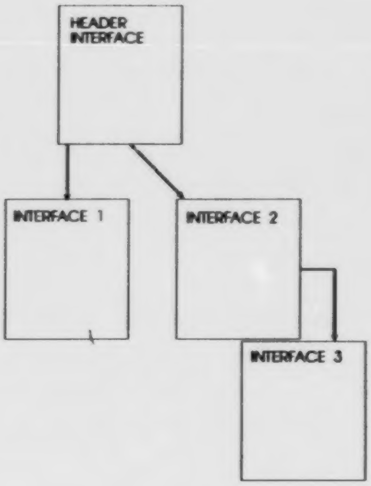

LAYOUT
ACTION

Interface

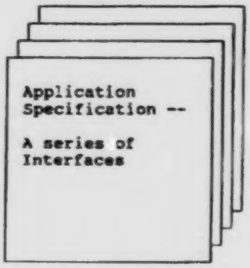
- used to define message(s) and prompts of an application
- consists of layout and action fields

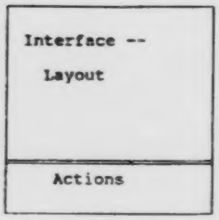
Application Specification

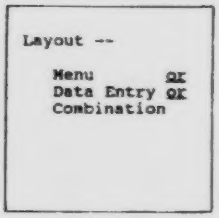
- used to handle an application
- collection of interfaces
- has a header interface


HEADER INTERFACE
INTERFACE 1
INTERFACE 2
INTERFACE 3

Application Specification

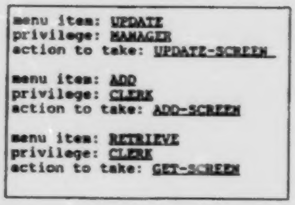

Application
Specification --
A series of
Interfaces


Interface --
Layout
Actions


Layout --
Menu or
Data Entry or
Combination

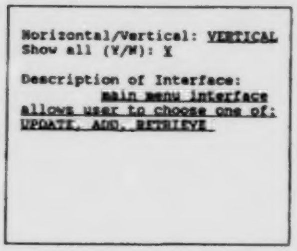
Example of the MENU LAYOUT DESIGNER

Enter Interface Name: MAIN-MENU
Type of layout: MENU



```
menu item: UPDATE
privilege: MANAGER
action to take: UPDATE-SCREEN

menu item: ADD
privilege: CLERK
action to take: ADD-SCREEN

menu item: RETRIEVE
privilege: CLERK
action to take: GET-SCREEN
```

<Enter>

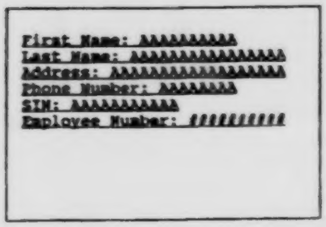


```
Horizontal/Vertical: VERTICAL
Show all (Y/N): Y

Description of Interface:
          main menu interface
allows user to choose one of:
UPDATE, ADD, RETRIEVE
```

<ENTER>

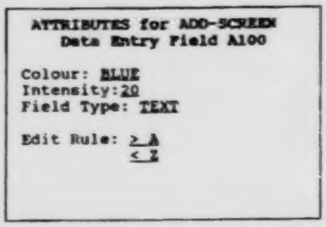
Example of the DATA ENTRY LAYOUT DESIGNER

Enter Interface name: ADD-SCREEN
Type of layout: DATA-ENTRY



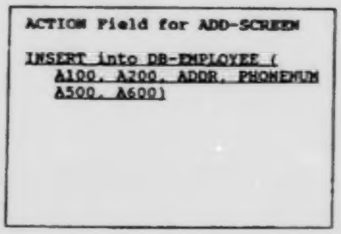
```
First Name: AAAAAAAAAA
Last Name: AAAAAAAAAAAAAAAA
Address: AAAAAAAAAAAAAAAAAA
Phone Number: AAAAAAAA
SIN: AAAAAAAAAAA
Employee Number: ##########
```

<ENTER>



```
  ATTRIBUTES for ADD-SCREEN
     Data Entry Field A100

Colour: BLUE
Intensity:20
Field Type: TEXT

Edit Rule: > A
           < Z
```

<ENTER>



```
ACTION Field for ADD-SCREEN

INSERT into DB-EMPLOYEE (
    A100, A200, ADDR, PHONENUM
    A500, A600)
```

<ENTER>

Example where a developer creates an Application Specification consisting of these interfaces

```
+--- Main Menu ----------+      +------- Reports --------+
|                        |      |                        |
|  Add Record            |      |    Report 1            |
|  Delete Record         |      |    Report 2            |
|  Reports               |      |    Report 3            |
|                        |      |                        |
|========================|      |========================|
| Call Add               |      | Call Report1           |
| Call Delete            |      | Call Report2           |
| Call Reports           |      | Call Report3           |
+------------------------+      +------------------------+
```

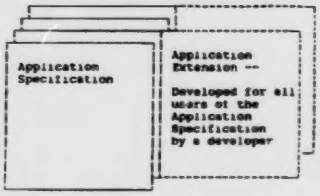
Developer at user site creates an Application Extension to create an Update Function consisting of deleting and adding a record. The Application Extension applies for all users using the application.

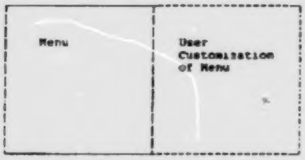
```
+--- Main Menu ----------+---- Main Menu ----+     +------ Update -----+
|                        |                   |     |                   |
|  Add Record            |  Update Record    |     |                   |
|  Delete Record         |  Reports          |     |                   |
|  Reports               |                   |     |                   |
|                        |                   |     |                   |
|========================|===================|     |===================|
| Call Add               | Call Update       |     | Call Delete       |
| Call Delete            | Call Reports      |     | Call Add          |
| Call Reports           |                   |     |                   |
+------------------------+-------------------+     +-------------------+
```

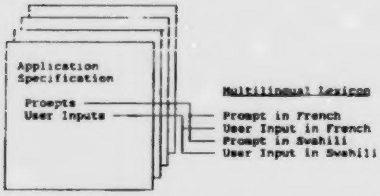
User Smith groups the two reports together and creates a Customized Menu for his own personal use.

```
+--- Reports ------------+----- Reports -----+     +------- EOM -------+
|                        |                   |     |                   |
|  Report 1              |  Report 1         |     |  Report 2         |
|  Report 2              |  End of Month     |     |  Report 3         |
|  Report 3              |                   |     |                   |
|                        |                   |     |                   |
|========================|===================|     |===================|
| Call Report1           | Call Report1      |     | Call Report2      |
| Call Report2           | Call EOM          |     | Call Report3      |
| Call Report3           |                   |     |                   |
+------------------------+-------------------+     +-------------------+
```

Multilingual Lexicon


Application Specification
Application Extension --
Developed for all users of the Application Specification by a developer


Menu
User Customization of Menu


Application Specification
Prompts
User Inputs
Multilingual Lexicon
Prompt in French
User Input in French
Prompt in Swahili
User Input in Swahili

A multilingual lexicon could exist for the application containing the language equivalents for any or all of the menu items.

Original language	French	Old English
Add record	Ajouter record	Insert record
Delete record		
Update record		
Report 1		
Report 2		
Report 3		
End of month	La fin du mois	Ye endeth of month

If an application specification is available in more than one language, it will run in the user's preferred language (if one exists) as set up in the user's profile. If the user's preferred language or the application specification has not been translated to the preferred language, then the default language for the application specification (as specified when the application specification is configured) will be used. If neither a user preferred language nor an application default language exists, the user will be prompted for the language desired if more than one language exists for the application. The user will also have the ability to switch languages while running an application specification.

Editor's note: The information contained in this discussion paper may not reflect the state of the Version H design at this time.

APPLICATION AND POPULARIZATION OF MINISIS IN CHINA

By Jiang Xiangdong
(MINISIS Resource Centre in China)

MINISIS entered China eight years ago. Under the auspices of IDRC, three organizations installed MINISIS software, and some users have done outstanding jobs with MINISIS. MINISIS has brought into full play in the information retrieval, library automation, and enterprise management.

1 How was MINISIS introduced to China?

I remember clearly, eight year ago, three IDRC experts, Miss Mary Campbell, Mr. Terras Gavin, and Mr. Richard Lee, who are the main designers and also at this meeting right now, came to Beijing, passed on MINISIS knowledge, installed MINISIS software for five new users, my own organization, STIC, among them.

The computerized information Retrieval started in China in 1974. The study, design and test of information storage and retrieval system of STIC started in 1975 and achieved with primary success in 1976. We installed bibliographic tapes of INSPEC, ISMEC, COMPENDEX and METADEX. In 1981, we obtained an HP3000/III computer from the Economic Assistance Project of UNDP. To create databases on the HP computer and provide users with information service, we needed software. An American consulting firm recommended MINISIS to us. In the autumn of 1981, three experts with MINISIS software stepped into the gate of China.

Before MINISIS, we used our own software to provide users with SDI information services, but we were not able to create online databases hence couldn't provide online searching services. MINISIS changed the state, it is an advanced tool for creating large databases and providing information services on HP computer. In the past eight years, MINISIS users in China have increased by a factor of five, from five in 1981 to thirty-three by now. MINISIS outstripped other information retrieval system software in terms of growth in user number in China.

The thirty-three users are located in nine cities and provinces, among them eighteen in Beijing and six in Shanghai. If they are classified by nature, thirteen are information centres of Ministry or Commission level, seven universities and colleges, five companies, three computer centres, two research institutes, two hospitals and one large factory. Judging from this, MINISIS has borne plentiful and substantial fruits in China. We sincerely thank IDRC and all MINISIS experts of IDRC for their tireless efforts.

2 MINISIS is important tool for information processing and library management in STIC.

STIC is among the first pioneer in computerized information retrieval in China. Fifteen years ago, we wrote retrieval programs by machine language on the computer made in China, and made a successful experiment in 1976, then we wrote SDI programs by COBOL on B6810 computer and HP 3000 computer and used them to search four bibliographic tapes, but those programs are not DBMS. By practice, we realized that it's a very difficult job to make a DBMS or IMS such as MINISIS. So, we decided to introduce MINISIS in 1980. It was a decision, I was very happy, from then on, I has been closely linked to MINISIS.

I attended first introductory MINISIS training course in Beijing in 1981, then I attended advanced MINISIS training courses in Philippines, Singapore and Ottawa under IDRC grants in the past four years. I am very grateful to Mr. Gavin, Mr. Lee and Miss Mary, they are all my respect teachers.

In the past eight years, we have tentatively established INSPEC, ISMEC, COMPENDEX, METADEX, JICST, and EMA bibliographic databases and Chinese character abstract database, and developed integrated library management system by use of MINISIS. We won some prizes from the State Scientific Commission, the State Economic Commission and the Ministry of Machinery and Electronics Industry for those achievements.

3 Establishment of MINISIS Resource Centre in China

In the past eight years, many persons worked with MINISIS in STIC, I had spent about 90% of my time on MINISIS. Since the spring, 1986, STIC has assisted IDRC in training Chinese users and popularized MINISIS in China. In 1986 and 1987, we have given MINISIS training courses to twenty-one users, and introduced MINISIS at six meetings to more than six hundred persons from three hundred organizations in the Beijing, Shanghai, Xian and Kunming. We publicized MINISIS at more than ten nation wide learning meetings and publishing papers on several magazines, so, MINISIS have already been widely known in the sphere of information retrieval, libraries, data processing and universities in China.

In the autumn, 1987, STIC established MRC under the support of IDRC, facts have proved that it is good idea to establish local MRC, because MRC can conveniently and timely provide services of training and technical supports. Because no language barriers exist between instructor and trainees, training is efficient, results is good.

In 1988 and 1989, MRC had done following works:

1)Established Chinese MINISIS User Association (CMUA)
In order to establish CMUA, MRC organized a meeting on October, 1988 in Beijing, more than 40 representatives from 24 MINISIS Users attended the meeting. Representatives discussed and worked out constitution of CMUA, and elected president, Vice presidents and secretary-general of CMUA. CMUA invited Mr. Richard Lee to assume honourary advisor. CMUA may be the first DBMS professional association in China.

2)Organized First Annual Meeting of CMUA.
The first annual meeting of CMUA was held in October, 1988 in Beijing, 45 persons from 24 users attended, Mr. Richard Lee was invited to introduce the new features of MINISIS version G, 9 persons of five users read their papers. The second annual meeting of CMUA will be held on Friday this week, Chinese representatives who attended MUG will attend the meeting.

3) Edited and published three issues of MINISIS newsletter for Chinese language users.
4) Wrote and published two books for MINISIS training.
Since 1982, Chinese users had been looking for MINISIS training documentation in Chinese language, to meet the needs, MRC published two books in 1988, one was MINISIS Bibliographic Database Management System for introductory MINISIS training, other was Interface Program Design on MINISIS Version G.

5) Provided MINISIS training courses.
In the recent two years, MRC trained 110 persons from 15 new users and several old users with introductory and advanced MINISIS training courses in Beijing,

Shanghai, Fujian Province, Henan Province, and Sichuan Province.

6) Translated and Published MINISIS Documentation.
A set of MINISIS Documentation including User's Manuals, Database Manager's Guide and Application Programmer's Guide, altogether 970 pages (about one and half million Chinese Characters) have been translated and published the last month.

7) Installed MINISIS software for new users and provided all users with technical supports.
In the recent two years, China MRC installed MINISIS for 15 new users and some old users, and helped users in resolving problems occurred during using MINISIS. Most problems were resolved by telephone, others by exchanging letters or by sending MRC personnel to users. The relationship between MRC and users is good. We like to give users any helps as far as possible, and share our knowledge and experience with all Chinese users. In order to popularize MINISIS in China, we have made efforts, and will make every effort.

4. The state of MINISIS applications in China.

The majority of HP 3000 computer users have become MINISIS users. About 70% of MINISIS users are new users starting later than 1986. Most users have brought in good results or made encouraging progress.

MINISIS applications in China concentrate in the following three aspects:

1) Library Management systems
MINISIS is very good DBMS for library management system, about ten users have created bibliographic data bases. The PAO SUI-LOONG Library of Shanghai Jiao Tong University have obtained gratifying achievements in this aspects, this library has developed a Multi-User Library Integrated System based on MINISIS and IMAGE, this system made full use powerful functions of Library management of MINISIS. Our library has developed integrated library management system also. Other organizations, for example, China Population information Centre, the Library of Chinese Academy of Medical Sciences, the Library of Shanghai Air Force Political College and the Library of China-Japan Friendship Hospital, have developed cataloguing subsystem, the acquisition subsystem and the periodical publications management system.

2) Create English and Chinese character databases.
Six users have created English bibliographic databases, they are: ISMEC, INSPEC, COMPENDEX, METADEX, JICST, EMA (my organization), CAB, AGRIS, ACRICOLA (Documentation Centre of Chinese Academy of Agricultural Science), API

(Luoyang Petrochemical Engineering Corporation), (Electronic Information Centre of State Sport Commission of China), SNOMED (China-Japan Friendship Hospital) and FGRAS (Asian pacific Regional Research and Training Centre for Integrated Fish Farming).

The Chinese character database are: Agricultural abstract database (Documentation Centre of Chinese Academy of Agricultural Science), Chinese Machinery Engineering abstract data base (my organization), Instrument Product database (Instrument Research Institute of the Ministry of Machinery and Electronics Industry). MINISIS Chinese interface provided users with basic conditions and complete functions for creating Chinese character database. Why did the Chinese users like to use MINISIS? There are three main reasons: first, it is powerful and advanced system: second, it can process Chinese characters; third, user can get it freely, and can obtain much supports from IDRC and MRC. I believe that more users will use MINISIS to create more Chinese character databases.

3) Develop Management Information System.
China-Japan Friendship Hospital has developed Ultrasonic Diagnosis Management System by use of MINISIS, this system had following functions: financial management, staff management, quality control and medical certificate management. The instrument product database actually is one kind of management information system. In China, a Hospital has developed management system by use of IMAGE before installing MINISIS, a large factory just installed MINISIS in the spring this year, I believe related users will gain more successful experiences in using MINISIS to manage enterprises from now on. MRC would like to help users to develop enterprise management system. I think that it is possible by means of MINISIS to develop a good enterprise management system.

By the way, several users have gained successful experiences in data transfer between MINISIS and other systems. Many Chinese MINISIS users have microcomputers, and have used CDS/ISIS or dBASE and created data bases. Some of them have resolved problems of data transfer between MINISIS and CDS/ISIS or dBASE. Shanghai Jiao Tong University has resolved problem of data transfer from CD-ROM to MINISIS.

Several users have used MINISIS extended application functions, for example, Jiao Tong University and China Population Information Centre have used MINISIS MARC interface, Documentation Centre of Chinese Academy of Agricultural Science have been providing some organizations with SDI services of CAB and AGRIS databases. China Population Information Centre and Shanghai Population Information Centre have succeeded in creating and using multilingual thesaurus.

Conclusion

From this time forth, quantity of MINISIS user will continuously increase, MRC will work hard and provide new users with good services of technical training and documentation supply, we will continue to help users to use MINISIS and continue to popularize MINISIS in China.

MINISIS is valuable resource, Let's share it to the best of one's ability.

MINISIS 在中国推广应用的八年回顾

机电部机械科技情报研究所　江向东
The Scientech Information Centre, MMEI
Jiang Xiangdong

【内容提要】 MINISIS 于 1981 年传入中国．目前我国已有三十多个 MINISIS 用户．本文回顾了八年来机械电子工业部科学技术情报所和我国其它一些用户应用 MINISIS 的情况．文章说．在过去几年中．机电部情报所使用 MINISIS 建立了 INSPEC, ISMEC, COMPENDEX, METADEX, JICST, EMA 等大型西文数据库和中国机械工程文摘数据库．开发了图书馆综合管理系统．其它一些中国用户在应用 MINISIS 方面也做了出色的工作．取得了可喜的成绩．同时．本文还顺便介绍了由加拿大国际开发研究中心支持的中国 MINISIS 技术研究中心在过去两年中究成的主要工作．

【Abstract】 MINISIS was introduced into China in 1981,there are more than thirty MINISIS Users in China now. This paper looks back to MINISIS applications in Sientech Information Centre (STIC) of Ministry of Machinery and Electronics Industry and other users of China in the past sevaral years.This paper said, STIC set up data bases of ISMEC, COMPENDEX, METADEX, INSPEC, JICST, EMA and Chinese Mechanical Engineering Abstract (CMEA), developed Integrated Library Management System by use of MINISIS, some user also done outstanding jobs on MINISIS with gratifying achievements. By the way, it also introduce the main works which were done by MRC Supported by IDRC in the past two years.

MINISIS 进入我国已整整八年了，八年来，在 IDRC 情报科学部的大力支持下，我国已有三十多个单位安装了 MINISIS，许多单位在应用 MINISIS 方面做了出色的工作，取得了可喜的成果．MINISIS 在我国的情报检索、图书馆自动化和各项管理工作中已经发挥了重要作用，而且必将发挥更大的作用．

一、MINISIS 是如何传入中国的

我国的计算机情报检索事业起步于 1974 年．我们单位从 1975 年开始进行计算机情报存储和检索系统的设计、研制和试验工作，并于 1976 年取得成功．接着从国外引进了 INSPEC、ISMEC、COMPENDEX 和 METADEX 文献磁带，1979 年从联合国计划开发署(UNDP)对华援助项目中，我们获得 HP3000Ⅲ计算机一台．为了使用 HP 计算机建立文献数据库，为中国用户提供情报服务，必须获得一个好的软件系统．通过美国一家顾问公司的介绍，我们决定引进于 1978 年研制成功的 MINISIS．

1981 年 10 月，IDRC 的三位专家，Mary Campbell 小姐，MINISIS 的主要研制者 Terras Gavin 先生和 Richard Lee 先生，来到北京，向我国五个 MINISIS 新用户的三十

多个学员传播MINISIS技术，为我国首批的五家用户安装MINISIS软件，并协助我单位开发应用项目。

八年过去了，我国的MINISIS用户已从起初的五个增加到现在的三十几个，它们分布在12个省市中，其中北京有18家用户，上海有6家用户。从用户的工作性质分，有14家部委级情报所，9家大专院校，6个国营公司，3个计算中心，2家大型医院，2个研究所，一家大型工厂。由此可知，MINISIS这一粒现代高科技的种子，不仅在中国的大地上生了根，发了芽，而且枝繁叶茂，并结出了丰硕的果实。

二、MINISIS是机电部情报所实现情报处理和图书管理现代化的主要手段

在中国，我们单位是最早使用计算机进行情报检索试验和检索服务的单位之一，从1975年开始，我们就着手设计情报检索程序和系统了，一开始用机器语言写检索程序，并在国产机上进行检索试验，然后用COBOL语言写程序，使用HP3000计算机为用户提供四种文献磁带的定题服务。1980年我们决定使用MINISIS。

过去八年中，我们使用MINISIS建立了ISMEC, COMPENDEX, METADEX, INSPEC等大型西文文献数据库以及中国机械工程文摘汉字数据库，建立了中西文图书管理系统。我们所取得的这些成果，曾分别受到机械电子工业部，国家经委和国家科委的六次嘉奖。

八年来，IDRC给了我们许多帮助，曾先后19次派人到我单位讲课、访问和进行技术合作（不包括这次年会）。例如，李先生曾十次远渡重洋，造访北京，帮助我们和中国用户解决了许多问题，对汉化MINISIS作出了贡献。

三、中国MINISIS技术研究中心的建立和工作

我们协助IDRC培训中国用户是从1986年春开始的，在1986与1987这两年中，我们曾给26个MINISIS用户提供过培训服务，并在上海，西安，昆明和北京等地十多次向300多个单位的600多人作过MINISIS报告。同时，我们还以论文、小册子、图书等形式，在全国的多种学术会议上和多种期刊杂志上宣传NINISIS。

1987年底，IDRC与我单位合作建立了中国MINISIS技术研究中心（MRC）。MRC可为本国或本地区众多的用户方便地及时地提供MINISIS培训服务及技术支持服务，由于教师与学员之间没有语言上的障碍，培训效率比较高，培训效果也比较好。用户一旦发生问题，一个电话或一封信便可解决问题。

在1988年初至1989年9月，MRC完成了以下工作:

1）建立了中国MINISIS用户协会

为了建立中国MINISIS用户协会，我们于1988年10月在北京召开了一次会议，来自24个用户的40多名代表参加了会议。会上，代表们讨论并通过了中国MINISIS用户协会章程，选举产生了理事长，秘书长和副理事长。我们还邀请IDRC的李先生担任协会的名誉顾问。

2）召开了中国 MINISIS 用户协会年会

中国 MINISIS 用户协会第一届年会于 1988 年 10 月在北京召开，来自 24 个用户的 45 名代表参加了会议，IDRC 的李先生应邀在会上作了 MINISIS G 版本新特色的报告，来自 5 个单位的 9 名代表宣读了论文。第二届年会于 1989 年 10 月召开，出席 MUG89 的中国用户代表都参加了这届年会。

3）编辑出版了三期 MINISIS 中文用户通讯

4）编写并出版了两本培训教材

自 1982 年以来，中国 MINISIS 用户一直希望获得 MINISIS 中文培训教材，为了满足他们的要求，我们于 1988 年出版了两本书，一本书名叫《MINISIS 中西文文献数据库管理系统》，作为 MINISIS 初级培训班的教材；另一本书名叫《MINISIS G 版接口程序设计》，用作 MINISIS 高级培训班的教材。

5）为 28 个用户提供了初级培训和高级培训服务

在过去的两年中，我们曾先后在北京、上海、福建、河南和四川等地举办了多期初级培训班和一期高级培训班，参加培训班的学员有 110 人，他们来自 15 个新用户和几个老用户，参加高级培训班的学员共 40 多人，他们来自 22 个用户。

6）翻译出版了 MINISIS G 版全套资料

MINISIS 全套资料由三部分组成："用户手册"，"数据库经理指南"和"应用程序员指南"。经过一年半的努力，全套中文资料已经出版，它由三本书组成，共 900 多页，约 130 万字。

7）为新用户提供 MINISIS 软件安装服务，为所有用户提供技术支持服务

过去两年中我们为 15 个新用户和几个老用户安装了 F 版或 G 版本软件，并协助各用户解决 MINISIS 应用中出现的许多问题。有时，我们也主动访问用户。我们希望每个中国用户在 MINISIS 应用方面获得成功，愿意给任何用户提供力所能及的帮助，多年来，我们一直为在中国推广 MINISIS 而努力工作，今后仍将如此。

四、MINISIS 在中国的应用情况

中国的大多数 HP3000 计算机用户都已成了 MINISIS 用户。在 30 多个用户中有 70%左右的用户是最近三年使用 MINISIS 的，有的用户今年刚安装 MINISIS，刚接受培训。今年的新用户刚开始使用 MINISIS 建立中文或西文数据库，所以这次他们来不及介绍经验了。MINISIS 的 30 多个用户中，大多数取得了积极的成果。

十家用户的近 20 名代表将在本届年会上宣读他们的论文，有十多个应用项目将进行演示。中国用户通过这次机会，展示他们在应用 MINISIS 方面所做的工作。

MINISIS 在中国的应用主要集中在以下三个方面：

1）建立图书管理系统

MINISIS 是一个优良的图书管理系统，我国已有近十个用户已经或正在使用 MINISIS 建立中西文馆藏图书资料数据库，并用来进行图书的收集、编目和检索等各项管理工作。在这方面比较突出的单位有上海交通大学包兆龙图书馆，该单位在 MINISIS 和 IMAGE 基础上开发了多用户图书综合系统（Multi-User Library Integrated

System–MILIS)，从而充分发挥了MINISIS强有力的图书管理功能．我们单位也已建立了功能类似的系统．此外，中国人口情报中心，中国医学科学院图书馆，上海空军政治学院，中日友好医院等单位也分别使用MINISIS建立了西文图书数据库，建立了图书编目、收集和期刊管理系统．还有一些单位也在进行类似的工作．

2) 建立西文和汉字文献数据库

在我国，使用MINISIS建立英文文献数据库的用户有六个，除本单位建立了前面已提及的四个大型文献数据库以外，还有中国农业科学院文献中心的CAB, AGRICOLA和AGRIS，洛阳石化总公司的API，中国体委电子中心的体育文摘数据库，中日友好医院的SNOMED及无锡淡水渔业研究中心的FGRAS．

使用MINISIS建立的汉字型文献数据库有：农科院文献中心的农业文摘数据库，本单位的机械工程文摘数据库，机电部仪器仪表数据库．MINISIS汉字接口的开发为中国用户建立汉字型数据库提供了良好的条件和比较齐全的功能．MINISIS是汉化最成功的软件系统．最近三年，MINISIS用户之所以迅速增加，主要原因有三个：第一，它本身功能强大，性能先进；第二，它能处理汉字；第三，IDRC免费向中国用户提供，并由中国MRC协助向中国用户提供技术服务．毫无疑问，今后将有更多的中国用户使用它建立其它汉字数据库．

3) 建立管理信息系统

利用MINISIS建立企业内部使用的管理信息系统，中日友好医院已进行了尝试．该单位已经在MINISIS的基础上建立了超声波诊断管理系统，该系统可用于进行超声波诊断方面的财务管理，人员管理，质量控制和病案管理．此外，机电部仪器仪表研究所建立的仪器仪表产品数据库实际上也属于管理信息系统．

在我国的MINISIS用户中，还有一家医院在安装MINISIS之前已经使用IMAGE建立了管理信息系统，有一家工厂今年刚安装MINISIS，所以，如何利用MINISIS去管理企业，还有待有关用户自己努力，我们MRC乐意与用户进行合作．

顺便说一说，在其它软件系统与MINISIS的数据交换方面，有几个用户已取得了成功．在我国的MINISIS用户中，几乎都拥有多台微型机，而且大部分都使用CDI／ISIS和dBASE建立了数据库．有些用户已经解决了这两个微机系统与MINISIS之间的数据传送问题．上海交通大学包兆龙图书馆还解决了MINISIS与CD-ROM之间的数据传送问题．有几个用户已经使用MINISIS扩充应用功能，如交通大学和中国人口情报中心使用了MINISIS-MARC接口，中国农科院文献中心通过MINISIS的SDI功能向200多用户提供CAB和AGRIS的定题服务．在建立和使用MINISIS多语言词表方面，中国人口情报中心和上海人口情报中心取得了成功的经验．

五、结束语

在今后，中国的MINISIS用户还会增加．中国MINISIS技术研究中心一定要将IDRC委托我们的所有工作做好，一定要为新用户提供良好的技术培训和资料服务，协助用户更有成效地使用MINISIS，使它在中国获得进一步推广．

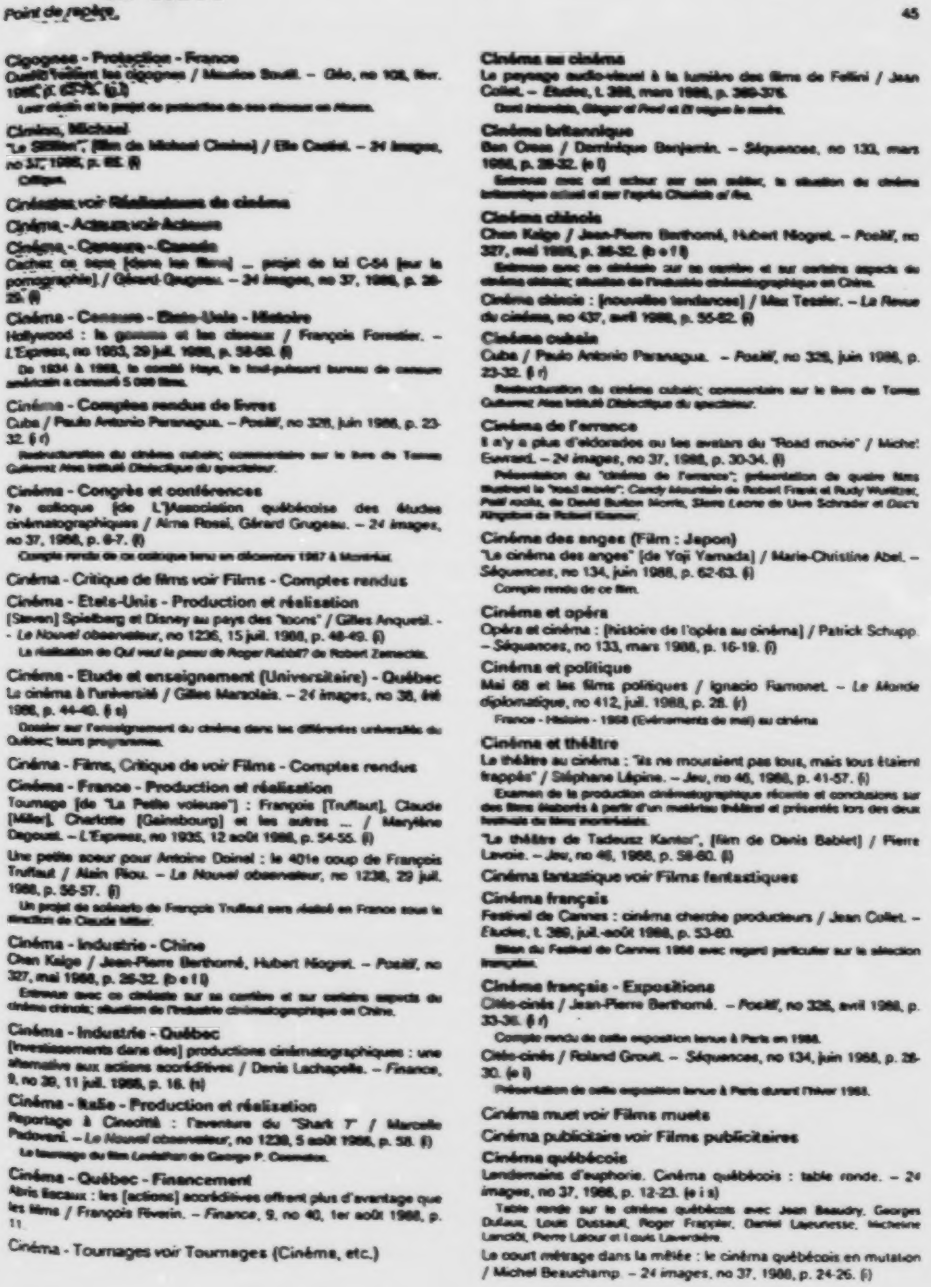
MINISIS at Services documentaires multimedia (SDM)

Alain Boucher
Services documentaires multimedia (SDM)

- Canada is a bilingual country (French, English.) The 6-million French-speaking community has its own information and library needs.
- A non-profit organization, Services documentaires multimedia (SDM) was established in 1964. Up to 1988, we were known as La Centrale des bibliothèques.
- Our aims: To develop, manage and disseminate bibliographic and textual databases in all fields of interest to users of French-language materials, in Canada and elsewhere.
- In addition, SDM is engaged in various cooperative projects with a wide range of organizations in Canada and other French-speaking countries (Europe, Africa.)
- Our staff: 95.
- Our budget: $C 4 000 000.
- We are a MINISIS user since 1981.
- Equipment:
    - HP 3000 series 58 (in process of upgrading to a Spectrum 950)
    - 35 terminals on site
    - 17 public access lines
    - 13 microcomputers (PC compatibles)
- Our databases:
    - A total of 50 databases
    - 1.2 million bibliographic records
    - 60,000 full text records (newspaper articles)
- Our own bibliographic databases:
    - 245,000 books
    - 140,000 articles from periodicals
    - 37,000 audiovisual materials
    - 3,000 software packages developed in Québec
    - 2,000 educational toys and plays
- In addition, 25 special interest databases developed in cooperation with, or managed at the request of various organizations, most of them of a public nature.
- Just to mention a few examples:
    - Educational research in Quebec
    - Newspaper articles from La Presse (major Montréal daily)
    - 225,000 bibliographic records
    - 60,000 full text articles
    - Geology and mineral resources of Québec
    - Research on leisure, recreation and tourism in French-speaking countries
    - Community health in Québec
    - Computer databases in Canada (French/English)

- All these databases are made available to libraries, public and private organizations, individuals, at moderate cost, in all desirable formats:

  printed publications
  library catalogue cards
  microfiche
  computer tapes and diskettes
  online access
  CD-ROM



Point de repère 45

Cigognes - Protection - France
[illegible] les cigognes / Maurice Soutif. – *Géo*, no 108, févr. 1988, p. [illegible]
Leur déclin et le projet de protection de ces oiseaux en Alsace.

Cimino, Michael
"Le Sicilien", [film de Michael Cimino] / Elie Castiel. – *24 images*, no 37, 1988, p. 65. (f)
Critique.

Cinéastes voir Réalisateurs de cinéma

Cinéma - Acteurs voir Acteurs

Cinéma - Censure - Canada
Cachez ce sexe [dans les films] ... projet de loi C-54 [sur la pornographie] / Gérard Grugeau. – *24 images*, no 37, 1988, p. 28-29. (f)

Cinéma - Censure - Etats-Unis - Histoire
Hollywood : la gomme et les ciseaux / François Forestier. – *L'Express*, no 1933, 29 juil. 1988, p. 58-59. (f)
De 1934 à 1968, le comité Hays, le tout-puissant bureau de censure américain a censuré 5 000 films.

Cinéma - Comptes rendus de livres
Cuba / Paulo Antonio Paranagua. – *Positif*, no 328, juin 1988, p. 23-32. (i r)
Restructuration du cinéma cubain; commentaire sur le livre de Tomas Gutierrez Alea intitulé *Dialectique du spectateur*.

Cinéma - Congrès et conférences
7e colloque [de L']Association québécoise des études cinématographiques / Alma Rossi, Gérard Grugeau. – *24 images*, no 37, 1988, p. 6-7. (f)
Compte rendu de ce colloque tenu en décembre 1987 à Montréal.

Cinéma - Critique de films voir Films - Comptes rendus

Cinéma - Etats-Unis - Production et réalisation
[Steven] Spielberg et Disney au pays des "toons" / Gilles Anquetil. -- *Le Nouvel observateur*, no 1236, 15 juil. 1988, p. 48-49. (f)
La réalisation de *Qui veut la peau de Roger Rabbit?* de Robert Zemeckis.

Cinéma - Etude et enseignement (Universitaire) - Québec
Le cinéma à l'université / Gilles Marsolais. – *24 images*, no 38, été 1988, p. 44-49. (i s)
Dossier sur l'enseignement du cinéma dans les différentes universités du Québec; leurs programmes.

Cinéma - Films, Critique de voir Films - Comptes rendus

Cinéma - France - Production et réalisation
Tournage [de "La Petite voleuse"] : François [Truffaut], Claude [Miller], Charlotte [Gainsbourg] et les autres ... / Marylène Dagouat. – *L'Express*, no 1935, 12 août 1988, p. 54-55. (i)
Une petite soeur pour Antoine Doinel : le 401e coup de François Truffaut / Alain Riou. – *Le Nouvel observateur*, no 1238, 29 juil. 1988, p. 56-57. (i)
Un projet de scénario de François Truffaut sera réalisé en France sous la direction de Claude Miller.

Cinéma - Industrie - Chine
Chen Kaige / Jean-Pierre Berthomé, Hubert Niogret. – *Positif*, no 327, mai 1988, p. 26-32. (b e f i)
Entrevue avec ce cinéaste sur sa carrière et sur certains aspects du cinéma chinois; situation de l'industrie cinématographique en Chine.

Cinéma - Industrie - Québec
[Investissements dans des] productions cinématographiques : une alternative aux actions accréditives / Denis Lachapelle. – *Finance*, 9, no 39, 11 juil. 1988, p. 16. (s)

Cinéma - Italie - Production et réalisation
Reportage à Cinecittà : l'aventure du "Shark T" / Marcelle Padovani. – *Le Nouvel observateur*, no 1239, 5 août 1988, p. 58. (i)
Le tournage du film *Leviathan* de George P. Cosmatos.

Cinéma - Québec - Financement
Abris fiscaux : les [actions] accréditives offrent plus d'avantage que les films / François Riverin. – *Finance*, 9, no 40, 1er août 1988, p. 11.

Cinéma - Tournages voir Tournages (Cinéma, etc.)

Cinéma au cinéma
Le paysage audio-visuel à la lumière des films de Fellini / Jean Collet. – *Etudes*, t. 368, mars 1988, p. 369-376.
Dont *Intervista*, *Ginger et Fred* et *Et vogue le navire*.

Cinéma britannique
Ben Cross / Dominique Benjamin. – *Séquences*, no 133, mars 1988, p. 28-32. (e i)
Entrevue avec cet acteur sur son métier, la situation du cinéma britannique actuel et sur l'après *Chariots of fire*.

Cinéma chinois
Chen Kaige / Jean-Pierre Berthomé, Hubert Niogret. – *Positif*, no 327, mai 1988, p. 26-32. (b e f i)
Entrevue avec ce cinéaste sur sa carrière et sur certains aspects du cinéma chinois; situation de l'industrie cinématographique en Chine.
Cinéma chinois : [nouvelles tendances] / Max Tessier. – *La Revue du cinéma*, no 437, avril 1988, p. 55-62. (i)

Cinéma cubain
Cuba / Paulo Antonio Paranagua. – *Positif*, no 328, juin 1988, p. 23-32. (i r)
Restructuration du cinéma cubain; commentaire sur le livre de Tomas Gutierrez Alea intitulé *Dialectique du spectateur*.

Cinéma de l'errance
Il n'y a plus d'eldorados ou les avatars du "Road movie" / Michel Euvrard. – *24 images*, no 37, 1988, p. 30-34. (i)
Présentation du "cinéma de l'errance"; présentation de quatre films illustrant le "road movie": *Candy Mountain* de Robert Frank et Rudy Wurlitzer, *Patti rocks*, de David Burton Morris, *Sierra Leone* de Uwe Schrader et *Doc's Kingdom* de Robert Kramer.

Cinéma des anges (Film : Japon)
"Le cinéma des anges" [de Yoji Yamada] / Marie-Christine Abel. – *Séquences*, no 134, juin 1988, p. 62-63. (i)
Compte rendu de ce film.

Cinéma et opéra
Opéra et cinéma : [histoire de l'opéra au cinéma] / Patrick Schupp. – *Séquences*, no 133, mars 1988, p. 16-19. (i)

Cinéma et politique
Mai 68 et les films politiques / Ignacio Ramonet. – *Le Monde diplomatique*, no 412, juil. 1988, p. 28. (r)
France - Histoire - 1968 (Evénements de mai) au cinéma

Cinéma et théâtre
Le théâtre au cinéma : "Ils ne mouraient pas tous, mais tous étaient frappés" / Stéphane Lépine. – *Jeu*, no 46, 1988, p. 41-57. (i)
Examen de la production cinématographique récente et conclusions sur des films élaborés à partir d'un matériau théâtral et présentés lors des deux festivals de films montréalais.
"Le théâtre de Tadeusz Kantor", [film de Denis Bablet] / Pierre Lavoie. – *Jeu*, no 46, 1988, p. 58-60. (i)

Cinéma fantastique voir Films fantastiques

Cinéma français
Festival de Cannes : cinéma cherche producteurs / Jean Collet. – *Etudes*, t. 369, juil.-août 1988, p. 53-60.
Bilan du Festival de Cannes 1988 avec regard particulier sur la sélection française.

Cinéma français - Expositions
Cités-cinés / Jean-Pierre Berthomé. – *Positif*, no 326, avril 1988, p. 33-36. (i r)
Compte rendu de cette exposition tenue à Paris en 1988.
Cités-cinés / Roland Groult. – *Séquences*, no 134, juin 1988, p. 28-30. (e i)
Présentation de cette exposition tenue à Paris durant l'hiver 1988.

Cinéma muet voir Films muets

Cinéma publicitaire voir Films publicitaires

Cinéma québécois
Lendemains d'euphorie. Cinéma québécois : table ronde. – *24 images*, no 37, 1988, p. 12-23. (e i s)
Table ronde sur le cinéma québécois avec Jean Beaudry, Georges Dufaux, Louis Dussault, Roger Frappier, Daniel Lajeunesse, Micheline Lanctôt, Pierre Latour et Louis Laverdière.
Le court métrage dans la mêlée : le cinéma québécois en mutation / Michel Beauchamp. – *24 images*, no 37, 1988, p. 24-26. (i)

Services documentaires multimédia

Ensemble du programme

001 DIAPOSITIVE

Dubé, Danièle

L'Ecologie / conception et réalisation de Danièle Dubé. – [Montréal] : Guérin, c1981.

133 diapositives : coul. + guide pédagogique (83 p.).

Titre du guide: Diaporama sur l'écologie québécoise.

Sommaire: Introduction à l'écologie à partir d'exemples québécois. Thèmes traités: notions générales, espèce et population, communauté, facteurs écologiques, écosystèmes, écosystème forestier, écosystème d'eau douce, écosystème marin.

ISBN 2-7601-0339-0 (guide)

*Sujet(s): *Ecologie - Etude et enseignement (Secondaire) - Québec (Province).*

574.5'07'09714/E17 SDM: 8320405
J* 5

Malgré le titre du guide, il ne s'agit pas d'un diaporama; tous les commentaires sont laissés au soin de l'enseignant. On trouve dans le guide pédagogique une introduction sur la méthodologie de l'ensemble et pour chaque diapositive: la description de son contenu, une question pour l'élève et des éléments de réponse.

*Descr. compl.: *Québec (Province) - Ecologie - Etude et enseignement (Secondaire)*

Distributeur(s): GUERIN

Module(s): 00

Module 1
Interrelations

002 FILM

Bêtes de proie / Nielsen-Ferns Ltd. – [Mississauga, Ont. : Marlin Motion Pictures], c1977.

1 bobine, env. 30 min. : son., coul. ; 16 mm. – (Une arche immobile)

Version anglaise produite en 1975 ss le titre: Predators. Générique: Producteur, Michael Maltby; réalisateurs, Len Gilday, Michael Maltby; scénario et narration, Gerald Durrell; musique, John Mills-Cockell; photographie, Len Gilday; montage, Steve Weslak.

Sommaire: Description de l'action et de l'importance de différents prédateurs. Regard sur certains de ces animaux vivant au Jersey Wildlife Preservation Trust.

*Sujet(s): *Animaux prédateurs *Jersey Zoological Park *Jardins zoologiques.*

591/U56 QL/758 SDM: 7821358
JA 5

Cette série d'information et de sensibilisation est passionnante. Le narrateur, directeur du zoo, connaît fort bien son sujet. Il s'adresse au spectateur avec naturel, chaleur et conviction, sur le ton d'une conversation à bâtons rompus et avec humour (on voit parfois des hommes à travers les cages). Il vulgarise avec aisance des notions complexes, les cas concrets abondent, et il dépasse souvent le niveau de la simple description pour aborder des problèmes écologiques; il apprend au spectateur à comprendre l'animal en fonction de ses propres caractéristiques et non à travers l'homme. Ainsi, l'animal insolite ou ordinaire devient toujours intéressant, sinon attachant. Les cadrages sont le plus souvent très rapprochés des animaux et les longueurs pratiquement absentes. L'ensemble est donc très riche et vivant.

Distributeur(s): MARLIN

Module(s): 01.05.01 - 01.05.02 - 04.05

003 FILM VIDEO

Biologie écologique / [consultant] J. Frank McCormick. – [Markham, Ont.] : Coronet Instructional Media ; [Loomis, Calif.] : Biomedia Associates, c1981.

1 bobine de film (16 min.) : son., coul. ; 16 mm. – (Sciences biologiques)

Titre original: Ecological biology. Offert également sur vidéo cassette.

Sommaire: Explique les diverses interrelations des organismes avec leur milieu et avec les autres êtres vivants, interrelations qui sont les fondements de l'écologie.

*Sujet(s): *Ecologie.*

*Sujet(s) jeunesse(s): *Ecologie.*

574.5/B615 SDM: 8521405
JA 4

*Descr. compl.: *Biologie - Ecologie*

Distributeur(s): CORONET

Module(s): 01.01

004 FILM

Le Boisé vivant. – Office national du film du Canada, 1960.

17 min. sonore. coul. 16 mm.

Version anglaise distribuée ss le titre: Life in the woodlot. Egalement disponible en 35 mm: 105C 0260 006. No de série: 106C 0260 006. Générique: Producteur, Hugh O'Connor; directeur, Tom Daly; réalisation et photographie, Dalton Muir; musique, Maurice Blackburn; montage, M. Russell.

Sommaire: Inventaire d'un boisé présentant les principaux spécimens des règnes animal et végétal qui y vivent; explication des lois internes qui régissent l'équilibre de ce milieu; conseils à l'exploitant sur la façon d'assurer la pérennité de son domaine.

*Sujet(s): *Forêts et sylviculture - Canada *Ecologie forestière - Canada.*

*Sujet(s) jeunesse(s): *Forêts *Ecologie *Nature.*

634.9'0971/B682 SDM: 7308871
E**JA 5

*Descr. compl.: *Boisés - Gestion *Canada - Forêts et sylviculture *Canada - Ecologie forestière*

Distributeur(s): ONF, RQ Licence: RQ

Module(s): 01.01

005 FILM

Les Carnivores. – C.T.V. Television Network/Metromedia Producers, [s.d.].

1 bobine (25 min.). – (Le Monde en liberté ; 121)

Offert également sur vidéo.

Sommaire: Descriptions et caractéristiques des carnivores, tels les loups, les lions les guépards et les hyènes. Explication du fait que ces créatures régularisent le cycle de la vie.

*Sujet(s): *Carnivores *Animaux prédateurs.*

599.74/C289 SDM: T200653

Distributeur(s): MAV Licence: SADA

Module(s): 01.05 - 03.07 - 04.05

006 FILM VIDEO

La Couleur dans la nature / version française, Multimedia audiovisuel en collab. avec Visual Education Centre. – [Chicago] : Encyclopaedia Britannica Educational Corporation, 1985, c1980.

12 min. : son., coul.

Accès sujet(s): *Archives - Gestion. *Documents administratifs - Gestion.

GD/30 SDM : 8827181

Ce logiciel de base de données textuelles permet la création, l'interrogation et l'extraction de compte-rendus, procès-verbaux et rapports. L'information versée dans la base de données peut être contrôlée à l'aide de validations. Le logiciel permet de produire un thésaurus et l'interrogation des données peut se faire par mots-clés ou dans le texte complet des documents.

Source: COBA

---

50 Q F

Gestion documentaire. – Brossard : Ordinateurs Hypocrat, [1986].

Pour ordinateurs sous systèmes d'exploitation MS-DOS (640K), UNIX ou XENIX (600K); disque rigide.

Pré-requis: système de gestion de bases de données INFORMIX de Relational Database Systems.

Prix sur demande

Accès sujet(s): *Documents administratifs - Gestion. *Archives - Gestion.

GD/30 SDM : 8727471

Destiné à tous les types d'entreprises, ce logiciel vise à rendre uniforme le classement des documents à en assurer le repérage rapide, à en permettre une épuration systématique et à restreindre au minimum le volume de la documentation.

Source: HYPOCRAT

---

51 Q F

Gestion documentaire Logi-plus. – Version 4.0. -- Verdun : Logibec Groupe Informatique, [1986].

Pour ordinateurs DEC VAX, IBM PC ou compatibles, 640K, VMS, MS-DOS/PC-DOS, UNIX, XENIX; disque rigide.

Pré-requis: version exécutable du logiciel Progress.

2800 $ (logiciel), y compris la formation [1988]

Accès sujet(s): *Documents administratifs - Gestion. *Archives - Gestion.

GD/30 SDM : 8627205

Conçu à partir du logiciel de base de données *Progress*, ce logiciel veut favoriser l'accessibilité à l'information au sein d'une entreprise. Le système fournit un classement normalisé des documents par unité administrative, catégorie, activité et sujet. L'accès aux documents se fait par recherche numérique ou alphabétique, par interrogation directe à l'écran ou via des index généraux ou spécialisés. Le système permet la gestion des délais de conservation des documents (actif, semi-actif, inactif), gère un plan de classification et effectue un déclassement automatique. Il offre de plus un contrôle de l'accès aux fonctions.

Source: LOGIBEC

---

52 Q F

Système de suivi de la correspondance Logi-plus. – Verdun : Logibec Groupe Informatique, [1987].

Pour ordinateurs sous systèmes d'exploitation UNIX, XENIX, VMS, MS-DOS, 640K (version DOS).

Prix sur demande

Accès sujet(s): *Correspondance - Gestion. *Archives - Gestion. *Documents administratifs - Gestion.

GD/30 SDM : 8727832

Ce logiciel facilite la gestion du courrier reçu par un organisme. Il permet: l'accès à différents fichiers de référence (répondants, codes géographiques); la gestion de la correspondance (information de base, synthèse du document, auteurs, signataires, descripteurs); le repérage de la correspondance; l'épuration de la correspondance; la production de listes énumératives, de rapports, de listes de rappel et de statistiques; l'impression d'un registre des transactions, etc. Le logiciel est disponible en mode mono-poste ou multi-usagers et son accès est contrôlé par mot de passe.

Source: LOGIBEC

---

53 Q F

Bibliographie plus / Gilles Fortier et Serge Berthelot. - - Montréal : Logidisque, [1988].

Pour ordinateurs Comterm MAX, IBM PC ou compatibles, 256K, MS-DOS/PC-DOS.

424,95 $ [1988]

Accès sujet(s): *Bibliographies. *Bases de données - Initiation. *Centres de documentation - Initiation. *Bibliothèques - Initiation.

GD/90 SDM : 8827739

Ce logiciel de base de données est dédié à la création et à la gestion de références bibliographiques organisées selon les normes reconnues (APA, ISO). L'usager peut créer une banque de descripteurs, entrer des bibliographies complètes (jusqu'à 200 titres par domaine de recherche), enregistrer des citations, des résumés ou des commentaires et retrouver ces informations selon divers modes de recherche. Il constitue également un instrument d'initiation à la gestion informatisée de bases de références bibliographiques, depuis l'entrée des données jusqu'à l'impression. Enfin, un tutoriel initie au fonctionnement des centres de documentation et à la consultation des ouvrages de référence.

Source: LOGIDISQUE

---

54 Q F

Le bibliophile. – Westmount : Servi-Comp, [1988].

Pour ordinateurs IBM PC ou compatibles, 650K, MS-DOS/PC-DOS; disque rigide, imprimante 132 colonnes.

1500 $ [1988]

Accès sujet(s): *Documents - Gestion.

GD/90 SDM : 8827777

Ce logiciel permet de gérer une collection de plusieurs milliers de documents: livres, périodiques, etc. Il permet d'effectuer des recherches par sujet et imprime diverses listes.

Source: SERVI-COMP

---

55 Q F

Dynamots. – Montréal : Logiciels Guépard, [1988].

Pour ordinateurs IBM PC ou compatibles, 640K, MS-DOS/PC-DOS.

249 $; 10 $ (démonstrateur) [1988]

Accès sujet(s): *Documents - Gestion.

GD/90 SDM : 8827582

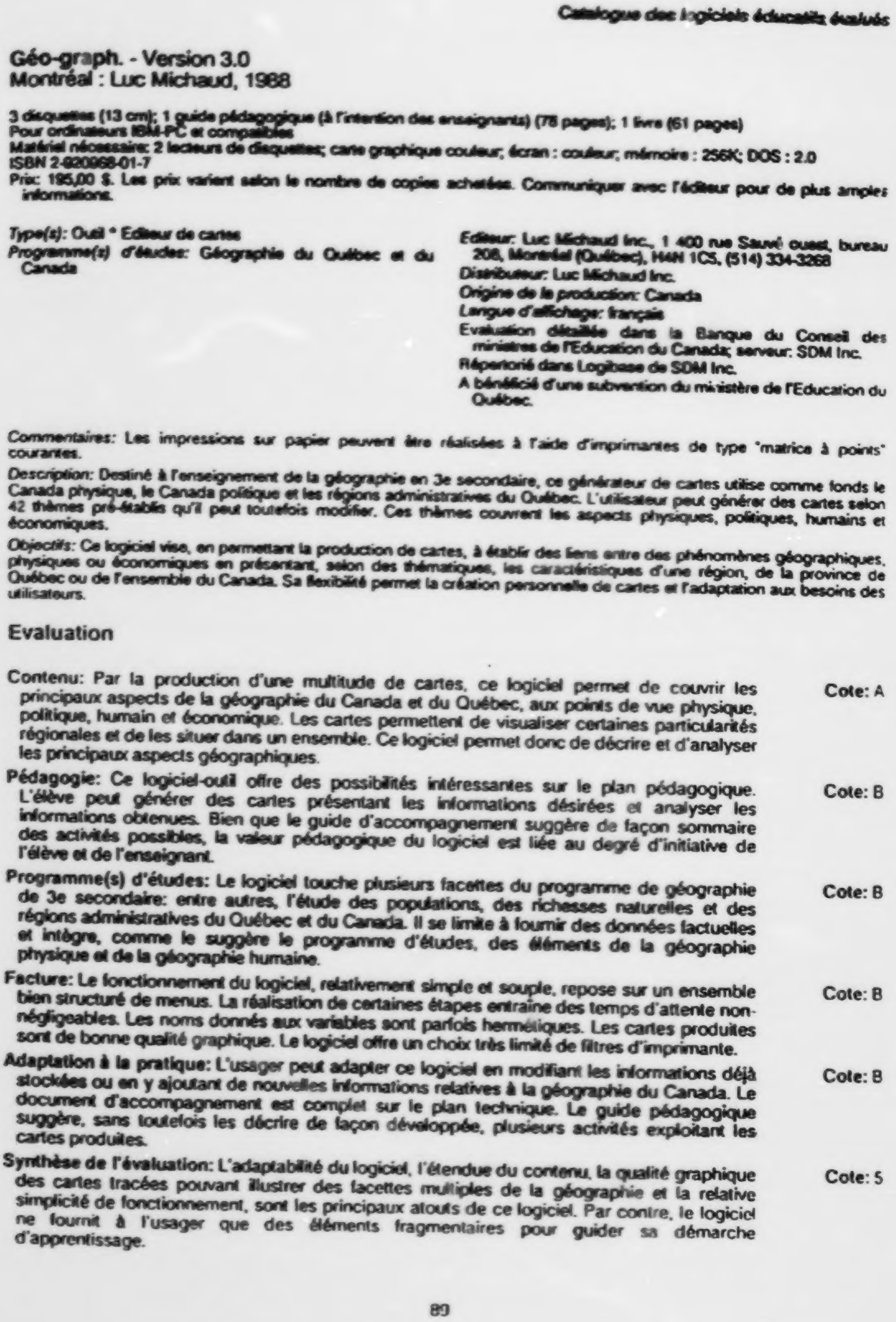
Ce progiciel permet de classer des collections d'objets ou de documents. La description des objets s'effectue en format libre et leur classification est contrôlée par un dictionnaire. Les recherches sont formulées en langage naturel et on peut recourir à la logique booléenne. Les objets repérés apparaissent à l'écran par ordre de pertinence et leur description peut être imprimée. Les écrans sont contrôlés par des fenêtres et l'aide y est contextuelle et modifiable.

Source: GUEPARD

---

56 Q F

Logidoc : gestion documentaire pour bibliothèque personnelle. -- Sherbrooke : Distributions évangéliques du Québec, [1988].



Géo-graph. - Version 3.0
Montréal : Luc Michaud, 1988

3 disquettes (13 cm); 1 guide pédagogique (à l'intention des enseignants) (78 pages); 1 livre (61 pages)
Pour ordinateurs IBM-PC et compatibles
Matériel nécessaire: 2 lecteurs de disquettes; carte graphique couleur; écran : couleur; mémoire : 256K; DOS : 2.0
ISBN 2-920968-01-7

Prix: 185,00 $. Les prix varient selon le nombre de copies achetées. Communiquer avec l'éditeur pour de plus amples informations.

Type(s): Outil * Editeur de cartes
Programme(s) d'études: Géographie du Québec et du Canada

Editeur: Luc Michaud Inc., 1 400 rue Sauvé ouest, bureau 208, Montréal (Québec), H4N 1C5, (514) 334-3268
Distributeur: Luc Michaud Inc.
Origine de la production: Canada
Langue d'affichage: français
Evaluation détaillée dans la Banque du Conseil des ministres de l'Education du Canada; serveur: SDM Inc.
Répertorié dans Logibase de SDM Inc.
A bénéficié d'une subvention du ministère de l'Education du Québec.

Commentaires: Les impressions sur papier peuvent être réalisées à l'aide d'imprimantes de type "matrice à points" courantes.

Description: Destiné à l'enseignement de la géographie en 3e secondaire, ce générateur de cartes utilise comme fonds le Canada physique, le Canada politique et les régions administratives du Québec. L'utilisateur peut générer des cartes selon 42 thèmes pré-établis qu'il peut toutefois modifier. Ces thèmes couvrent les aspects physiques, politiques, humains et économiques.

Objectifs: Ce logiciel vise, en permettant la production de cartes, à établir des liens entre des phénomènes géographiques, physiques ou économiques en présentant, selon des thématiques, les caractéristiques d'une région, de la province de Québec ou de l'ensemble du Canada. Sa flexibilité permet la création personnelle de cartes et l'adaptation aux besoins des utilisateurs.

Evaluation

Contenu: Par la production d'une multitude de cartes, ce logiciel permet de couvrir les principaux aspects de la géographie du Canada et du Québec, aux points de vue physique, politique, humain et économique. Les cartes permettent de visualiser certaines particularités régionales et de les situer dans un ensemble. Ce logiciel permet donc de décrire et d'analyser les principaux aspects géographiques. **Cote:** A

Pédagogie: Ce logiciel-outil offre des possibilités intéressantes sur le plan pédagogique. L'élève peut générer des cartes présentant les informations désirées et analyser les informations obtenues. Bien que le guide d'accompagnement suggère de façon sommaire des activités possibles, la valeur pédagogique du logiciel est liée au degré d'initiative de l'élève et de l'enseignant. **Cote:** B

Programme(s) d'études: Le logiciel touche plusieurs facettes du programme de géographie de 3e secondaire: entre autres, l'étude des populations, des richesses naturelles et des régions administratives du Québec et du Canada. Il se limite à fournir des données factuelles et intègre, comme le suggère le programme d'études, des éléments de la géographie physique et de la géographie humaine. **Cote:** B

Facture: Le fonctionnement du logiciel, relativement simple et souple, repose sur un ensemble bien structuré de menus. La réalisation de certaines étapes entraîne des temps d'attente non-négligeables. Les noms donnés aux variables sont parfois hermétiques. Les cartes produites sont de bonne qualité graphique. Le logiciel offre un choix très limité de filtres d'imprimante. **Cote:** B

Adaptation à la pratique: L'usager peut adapter ce logiciel en modifiant les informations déjà stockées ou en y ajoutant de nouvelles informations relatives à la géographie du Canada. Le document d'accompagnement est complet sur le plan technique. Le guide pédagogique suggère, sans toutefois les décrire de façon développée, plusieurs activités exploitant les cartes produites. **Cote:** B

Synthèse de l'évaluation: L'adaptabilité du logiciel, l'étendue du contenu, la qualité graphique des cartes tracées pouvant illustrer des facettes multiples de la géographie et la relative simplicité de fonctionnement, sont les principaux atouts de ce logiciel. Par contre, le logiciel ne fournit à l'usager que des éléments fragmentaires pour guider sa démarche d'apprentissage. **Cote:** 5

08157

Parlons avec Ploum = Ploum gets us talking / Christine Nicolas. — [Paris] : Nathan, c1985 ; [Montréal : Brault & Bouthillier, distributeur].

1 jeu (64 cartes-images, 14 morceaux de puzzles) : carton, coul. ; dans une boîte, 37 x 26 x 5 cm + feuillet pédagogique.

Pour 1-6 joueurs. Contrairement aux indications données sur la boîte, le nombre de pièces du casse-tête est de 14. Nathan : 337020.

Descripteurs ESAR: *Jeu de séquence

*Classification *Correspondance de regroupements *Relations temporelles

*Reproduction d'événements *Reproduction de rôles *Discrimination visuelle

*Activité associative

*Appellation verbale *Expression verbale *Mémoire phonétique

*Identification parentale

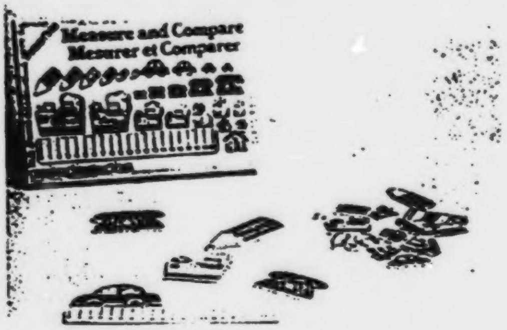
Descripteurs complémentaires: *Vie quotidienne *Ours

Analyse psychologique: Ensemble de soixante-quatre images cartonnées représentant Ploum l'ourson et sa famille dans des scènes de la vie quotidienne. Plusieurs effigies cartonnées de Ploum peuvent, de plus en plus, être assemblées: six visages aux mimiques différentes (sourire, larmes, moue, etc.), des corps et des jambes sont prévus à cet effet. L'enfant très jeune reconstitue ces casse-tête simples et nomme les vêtements et les parties du corps de l'ourson. Il observe et décrit ensuite les illustrations colorées, les classant selon des critères subjectifs (personnages, activités ou moments de la journée). Il assimile les notions de temps pour ordonner les activités de Ploum au cours d'une journée. Il exprime ses pensées et ses sentiments relatifs aux émotions vécues par le jeune héros auquel il s'identifie et plus tard apprivoise les caractéristiques particulières des autres membres de la famille. Ce jeu s'utilise en petit groupe d'enfants stimulés par la présence d'un adulte.

Analyse matérielle: Ploum, petite mascotte des jeux Nathan, permet ici d'aborder le thème des émotions vécues par les enfants dans leurs activités quotidiennes. Les illustrations simples et colorées contiennent juste assez de détails pour stimuler les discussions et laisser un peu de place à l'imagination. Fortement recommandé pour les collectivités: services de garde, classes maternelles et pour les cliniques de psychothérapie infantile.

Âge: 2 à 8 ans.
Évaluation: Très recommandé
Prix approximatif: 78,49 $

08158



Mesurer et comparer = Measure and compare. — Plattsburgh, N. Y. ; Montréal : Battat, [1988?].

1 jeu (règle, 25 illustrations) : bois, coul. ; dans une boîte, 19 x 28 x 4 cm + feuillet de règles. — (Battat games)

Pour 1-4 joueurs. Battat: 2021.

Descripteurs ESAR: *Jeu de séquence

*Sériation *Opérations numériques *Raisonnement concret

*Application de règles *Discrimination visuelle

*Activité solitaire ou compétitive

*Appellation verbale *Expression verbale *Mémoire phonétique

Descripteurs complémentaires: *Mesures

Analyse psychologique: L'enfant s'applique à sérier des objets en ordre croissant et décroissant. Par observation il constate les différentes dimensions et les compare. Si on joue en groupe, un meneur de jeu pose des questions en décrivant verbalement les différences de dimensions. Ce jeu amène l'enfant à poser des hypothèses et à discuter ou expliquer les approximations qu'il propose. Il faut répondre d'abord par simple observation et on mesure ensuite avec la règle en centimètres pour confirmer. Il nomme et identifie la longueur des objets. Ce jeu de sériation implique la participation de l'adulte pour enseigner les comparatifs et les superlatifs. On peut aussi l'utiliser pour se familiariser avec les chiffres et les unités de mesure. Jeu individuel ou compétitif.

Analyse matérielle: Jeu utilisé surtout dans un cadre pédagogique avec l'aide d'un adulte. Recommandé dans les services de garde, les maternelles et les classes du primaire.

Âge: 4 à 7 ans.
Évaluation: Recommandé
Prix approximatif: 12,99 $

ENVIRONNEMENT

182

CENV **

(Canadian Environment)

Notices bibliographiques et résumés de la documentation scientifique et technique relative à l'environnement canadien.

Type de données: bibliographiques
Langue(s): anglais
Date de départ: 1970
Taille: 65 000 enregistrements (1988-05)

Producteur: Environnement Canada, Direction générale des eaux et terres, WATDOC, Ottawa (Ontario) K1A 0E7
Contact(s): Louise Power, (819) 997-2324
Serveur(s): QL

183

DREF ****

(Data Reference System)

Répertoire des systèmes canadiens d'information dans le domaine des ressources hydriques, des terres, de la faune, de la foresterie et des autres sujets reliés à l'environnement. Les données proviennent des services d'Environnement Canada, d'organismes fédéraux et provinciaux, d'universités et d'organismes privés.

Type de données: bibliographiques
Langue(s): anglais
Date de départ: 1900

Producteur: Environnement Canada, Direction générale des eaux et terres, WATDOC, Ottawa (Ontario) K1A 0E7
Contact(s): Louise Power, (819) 997-2324
Serveur(s): QL

184

ELIAS **

(Environment Libraries Automated System)

Collections documentaires des bibliothèques participantes et d'Environnement Canada.

Type de données: bibliographiques
Langue(s): anglais
Date de départ: 1972
Taille: 47 000 enregistrements (1989-01)
Accroissement: 3 500 enregistrements par année

Producteur: Environnement Canada, Bibliothèque, Ottawa (Ontario) K1A 1C7
Serveur(s): CAN/OLE – iNet 2000

185

ENV ***

(Environnement)

Références bibliographiques à la documentation scientifique et technique de langue française relative à l'environnement canadien.

Type de données: bibliographiques
Langue(s): français
Date de départ: 1971
Taille: 12 000 enregistrements (1988-05)

Producteur: Environnement Canada, Direction générale des eaux et terres, WATDOC, Ottawa (Ontario)
Contact(s): Louise Power, (819) 997-2324
Serveur(s): QL

186

RSCPT/LEGAL = IRPTC/LEGAL **

(Registre international des substances chimiques potentiellement toxiques = International Register of Potentially Toxic Chemicals)

Texte intégral des lois et règlements régissant l'utilisation des produits toxiques au Canada. Comprend aussi des informations de nature juridique provenant de 22 autres nations ainsi que d'organismes internationaux.

Type de données: primaires
Langue(s): français et anglais
Date de départ: 1988

Producteur: Santé et Bien-être Canada, Tunny's Pasture, Ch. 212, Ottawa (Ontario) K1A 0L2
Contact(s): (613) 957-1506
Serveur(s): iNet 2000

La pollution par la fumée du tabac est négligeable dans les bureaux. -- (PC)

Réf.: 881001, p. A9.
Type: nouv (moyen)

Descripteurs: *CSN. *Sterling, Theodor D. *Pollution domestique. *Bâtiments malades, Syndrome des. *Pollution par la fumée de tabac.

QUÉBEC -- Quelque 500 enquêtes menées dans des édifices malsains et des centaines d'études sur les polluants ambiants montrent que l'atmosphère d'un édifice demeure hautement polluée, même en l'absence de fumeurs, s'il ne bénéficie pas d'une ventilation adequate et d'un entretien conforme.

C'est là la conclusion d'une conférence prononcée hier au colloque de la CSN sur les problèmes de la ventilation par le Dr Theodor D. Sterling, de la Faculté de sciences appliquées de l'université Simon Fraser, de Burnaby, en Colombie-Britannique.

Le colloque de la CSN qui prend fin aujourd'hui avait débuté jeudi et se déroulait sous le thème de «l'air malade... dans les bâtiments malsains». On a pris l'habitude de qualifier «d'édifices malsains» les édifices dont les occupants se plaignent de différents problèmes de santé. Et l'Organisation mondiale de la santé a regroupé l'ensemble de ces plaintes sous l'expression de «syndrome des tours à bureaux».

On a souvent avancé que la fumée de cigarette était responsable de la plupart des problèmes de qualité de l'air ambiant, en raison de sa visibilité. «Cependant, une multitude de faits démontrent le contraire», a déclaré le Dr Sterling.

Depuis 1971, l'Institut national américain de la santé et de la sécurité au travail a inspecté 484 édifices posant des problèmes de santé et de confort. Une analyse des résultats de 203 de ces inspections, terminées vers la fin de 1983, montre que la fumée de cigarette a été reconnue comme facteur responsable dans 2 p. cent seulement des cas étudiés, la ventilation inadéquate étant le problème le plus fréquent, c'est-à-dire dans 48,3 p. cent des cas.

L'analyse des résultats a aussi montré qu'une vaste gamme d'éléments contribue à la mauvaise qualité de l'air ambiant et n'a pas réussi à établir un lien significatif entre la fumée présente dans l'environnement et les problèmes de santé et de confort posés par ces édifices.

L'analyse a aussi montré qu'à l'exception des matériaux de construction dangereux comme l'amiante, la ventilation inadéquate constitue, de très loin, le problème majeur.

«La cause secondaire la plus grave semble être la contamination de l'alimentation en air frais par les émanations des véhicules extérieurs ou des restaurants voisins, ou encore, par la réintroduction de l'air vicié provenant de l'édifice.»

L'enquête, par ailleurs, démontre clairement que la suppression totale de la fumée présente dans l'environnement n'éliminerait pas totalement la plupart des substances contenues dans la fumée de cigarette.

- We are very happy with MINISIS and with IDRC support (which is seldom required!)
- Our development and interests:

  Full text applications

  Typeset quality publications with laser printers on microcomputers, from MINISIS databases

  CD-ROM

  User interface with MINISIS (user friendly communication software on microcomputers)
- Typeset quality publications

  15,000 master pages per year which are offset printed

  Direct use of output from the PRINT processor, with embedded typesetting codes

  Automatic downloading of data on microcomputer, during nights and week-ends

  Final processing and printing using Microsoft Word and HP Laserjet II printers

```
{\rtf1\pc
\gutter144\margl864\margr1008\margt1152\margb1080\ftnbj\ftnrestart\widowct
rl
\sectd \sbknone\linex576\headery576\cols2\colsx432\endnhere {\header
\pard \qc\keep\sl-180 \plain \b\f8\fs16 FACTUAL * Mise \'85 jour du 1er
septembre 1987 * page \chpgn \par}\pard\keep\li144\fi-144\sl-
180\sb30\plain\f8\fs16 Accident nucléaire de Tchernobyl, 1986
\plain\i\f8\fs16 voir \plain\b\f8\fs16 Tchernobyl, Accident nucléaire de,
URSS, 1986 \par \pard\keep\li144\fi-144\sl-180\sb30\plain\f8\fs16 Accord
de libre-échange canado-américain \plain\i\f8\fs16 voir \plain\b\f8\fs16
Libre-échange - Canada-Etats-Unis \par \pard\keep\li144\fi-144\sl-
180\sb30\plain\f8\fs16 Action bénévole \plain\i\f8\fs16 voir
\plain\b\f8\fs16 Bénévolat \par \pard\keep\li144\fi-144\sl-
180\sb30\plain\f8\fs16 Age mésolithique  \plain\i\f8\fs16 voir
\plain\b\f8\fs16 Mésolithique \par \pard\keep\li144\fi-144\sl-
180\sb30\plain\f8\fs16 Agence de protection de l'environnement (Etats-
Unis) \plain\i\f8\fs16 voir \plain\b\f8\fs16 Etats-Unis. Environmental
Protection Agency \par \pard\keep\li144\fi-144\sl-180\sb30\plain\f8\fs16
Aide bénévole  \plain\i\f8\fs16 voir \plain\b\f8\fs16 Bénévolat \par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Aliments congelés -
Industrie et commerce \plain\f8\fs16 [2 * i * s]\par \pard\keep\li288\sl-
170\plain\f8\fs16 x Industrie de la surcongélation\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Aliments congelés -
Industrie et commerce - Québec\par \pard\keep\li144\fi-144\sl-
180\sb30\plain\b\f8\fs16 Appel (Droit) \plain\f8\fs16 [42 * i * s]\par
\pard\keep\li288\sl-170\plain\f8\fs16 x Droit d'appel\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\f8\fs16 Ara-C \plain\i\f8\fs16
voir \plain\b\f8\fs16 Cytosine arabinoside \par \pard\keep\li144\fi-
144\sl-180\sb30\plain\f8\fs16 Arabes, Etats  \plain\i\f8\fs16 voir
\plain\b\f8\fs16 Etats arabes \par \pard\keep\li144\fi-144\sl-
180\sb30\plain\f8\fs16 Archevêchés  \plain\i\f8\fs16 voir \plain\b\f8\fs16
Diocèses \par \pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16
Architecture et handicapés \plain\f8\fs16 [70 * c * s]\par
\pard\keep\li288\sl-170\plain\f8\fs16 x Barrières architecturales\par x
Handicapés et architecture\par \pard\keep\li144\fi-144\sl-
180\sb30\plain\b\f8\fs16 Architecture et handicapés - France\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Association
internationale du transport aérien \plain\f8\fs16 [92 * i * c ]\par
\pard\keep\li288\sl-170\plain\f8\fs16 x IATA\par \pard\keep\li144\fi-
144\sl-180\sb30\plain\f8\fs16 Balouba (Peuple d'Afrique) \plain\i\f8\fs16
voir \plain\b\f8\fs16 Baluba (Peuple d'Afrique) \par \pard\keep\li144\fi-
144\sl-180\sb30\plain\b\f8\fs16 Baluba (Peuple d'Afrique) \plain\f8\fs16
[96 * c * c]\par \pard\keep\li288\sl-170\plain\f8\fs16 x Balouba (Peuple
d'Afrique)\par x Louba (Peuple d'Afrique)\par x Luba (Peuple
d'Afrique)\par \pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Baluba
(Peuple d'Afrique) - Folklore\par \pard\keep\li144\fi-144\sl-
180\sb30\plain\f8\fs16  Barrières architecturales \plain\i\f8\fs16 voir
\plain\b\f8\fs16 Architecture et handicapés \par \pard\keep\li144\fi-
144\sl-180\sb30\plain\f8\fs16 Bateaux de plaisance - Réfrigération
\plain\i\f8\fs16 voir \plain\b\f8\fs16 Réfrigération à bord des bateaux de
plaisance \par \pard\keep\li144\fi-144\sl-180\sb30\plain\f8\fs16 Bébés -
Mort subite  \plain\i\f8\fs16 voir \plain\b\f8\fs16 Mort subite chez le
nourrisson \par \pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16
Bénévolat \plain\f8\fs16 [59 * c * s]\par \pard\keep\li288\sl-
170\plain\f8\fs16 x Action bénévole\par x Aide bénévole\par x
Collaboration bénévole\par x Service bénévole\par x Travail bénévole\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Bénévolat - Canada\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Bénévolat - France\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Bénévolat - Québec\par
\pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Bénévolat - Québec -
Beauport\par \pard\keep\li144\fi-144\sl-180\sb30\plain\b\f8\fs16 Bénévolat
```

FACTUAL * Mise à jour du 1er septembre 1988 * page 1

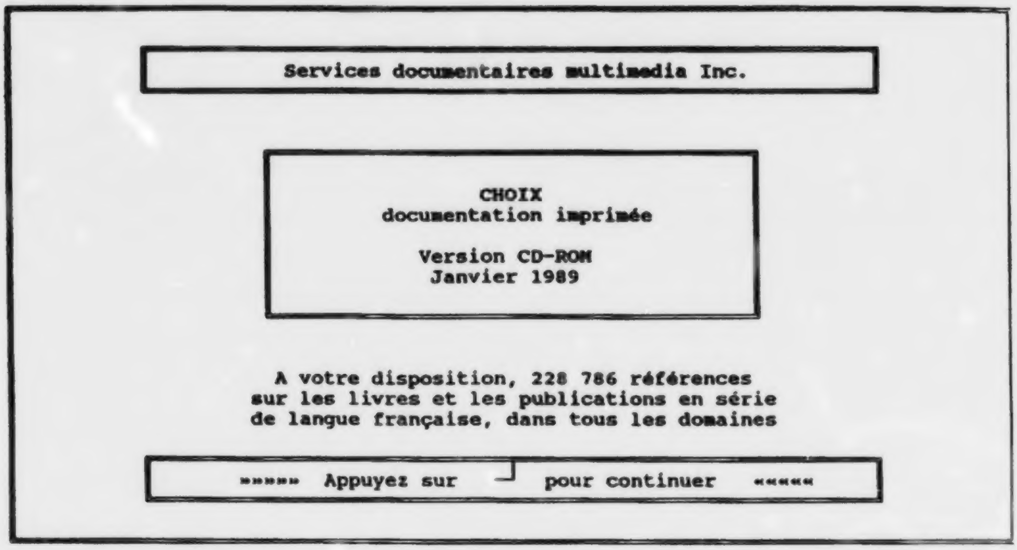
Accident nucléaire de Tchernobyl, 1986 *voir* Tchernobyl, Accident nucléaire de, URSS, 1986
Accord de libre-échange canado-américain *voir* Libre-échange - Canada-Etats-Unis
Action bénévole *voir* Bénévolat
Age mésolithique *voir* Mésolithique
Agence de protection de l'environnement (Etats-Unis) *voir* Etats-Unis. Environmental Protection Agency
Aide bénévole *voir* Bénévolat
Aliments congelés - Industrie et commerce [2 * l * s]
    x Industrie de la surcongélation
Aliments congelés - Industrie et commerce - Québec
Appel (Droit) [42 * l * s]
    x Droit d'appel
Ara-C *voir* Cytosine arabinoside
Arabes, Etats *voir* Etats arabes
Archevêchés *voir* Diocèses
Architecture et handicapés [70 * c * s]
    x Barrières architecturales
    x Handicapés et architecture
Architecture et handicapés - France
Association internationale du transport aérien [92 * l * c]
    x IATA
Balouba (Peuple d'Afrique) *voir* Baluba (Peuple d'Afrique)
Baluba (Peuple d'Afrique) [96 * c * c]
    x Balouba (Peuple d'Afrique)
    x Louba (Peuple d'Afrique)
    x Luba (Peuple d'Afrique)
Baluba (Peuple d'Afrique) - Folklore
Barrières architecturales *voir* Architecture et handicapés
Bateaux de plaisance - Réfrigération *voir* Réfrigération à bord des bateaux de plaisance
Bébés - Mort subite *voir* Mort subite chez le nourrisson
Bénévolat [59 * c * s]
    x Action bénévole
    x Aide bénévole
    x Collaboration bénévole
    x Service bénévole
    x Travail bénévole
Bénévolat - Canada
Bénévolat - France
Bénévolat - Québec
Bénévolat - Québec - Beauport
Bénévolat - Québec - Saint-Basile-le-Grand
Bibliothèques - Services aux handicapés *voir* Bibliothèques et handicapés
Bibliothèques et handicapés [71 * c * s]
    x Bibliothèques - Services aux handicapés
    x Handicapés et bibliothèques
Binandele (Peuple papou) [97 * c * c]
    x Orokaiva (Peuple papou)
Binandele (Peuple papou) - Moeurs et coutumes
Bio-industries [94 * d * s]
    x Bioindustries
Bioindustries *voir* Bio-industries
Blaireau américain [74 * l * s]
    x Blaireau d'Amérique
    x Carcajou
Blaireau américain - Moeurs et comportement
Blaireau d'Amérique *voir* Blaireau américain
Botanique - Histologie *voir* Plantes - Cellules et tissus
Camions [65 * c * s]
    x Poids lourds (Camions)
Canal 10 (Station de télévision) *voir* Télé-Métropole (Station de télévision)
Cancer chez l'enfant [46 * d * s]
    x Cancérologie pédiatrique
    x Tumeurs chez l'enfant
Cancer chez l'enfant - Aspect psychologique
Cancer chez l'enfant - Diagnostic
Cancer chez l'enfant - Traitement
Cancérologie pédiatrique *voir* Cancer chez l'enfant
CAO *voir* Conception assistée par ordinateur
Carcajou *voir* Blaireau américain
CD - Vidéo *voir* Vidéodisques compacts
Cégeps [21 * d * s]
    x Collèges d'enseignement général et professionnel
    Pour les cégeps particuliers voir au nom officiel. Ex. Cégep du Vieux-Montréal; Cégep de Jonquière
Cégeps - Pastorale
Cellules végétales *voir* Plantes - Cellules et tissus
Centres dramatiques *voir* Théâtres
Cerveau - Hémorragie [47 * c * s]
    x Hémorragie cérébrale
Cerveau - Hémorragie - Traitement
CFTM (Station de télévision) *voir* Télé-Métropole (Station de télévision)
Championnats internationaux de tennis *voir* U.S. Open (Tennis professionnel : New York, Etats-Unis)
Charan industries (Firme) *voir* Industries Charan (Firme)
Chaussées *voir* Revêtements (Voirie)
Chemins *voir* Routes
Chênes [82 * c * s]
    x Quercus
Chimie appliquée *voir* Chimie industrielle
Chimie industrielle [66 * c * s]
    x Chimie appliquée
    x Génie chimique
    x Technologie chimique
Chimie industrielle - Aspect de l'environnement
Chimie industrielle - Sécurité - Mesures
Cinéma muet *voir* Films muets
Cliniques de sang *voir* Collecte de sang
Coeur - Rejet [48 * l * s]
    x Rejet cardiaque
Collaboration bénévole *voir* Bénévolat
Collecte de sang [5 * l * s]
    x Cliniques de sang
    x Sang, Collectes de
Collecte de sang - Québec
Collecte de sang - Québec - Montréal
Collèges d'enseignement général et professionnel *voir* Cégeps
Cols bleus *voir* Employés municipaux
Compact disc vidéo *voir* Vidéodisques compacts
Compagnie pétrolière impériale (Firme) [25 * l * c]
    x Imperial Oil (Firme)
    x Impériale (Firme)
    x Pétrolière impériale (Firme)
Conception assistée par ordinateur [23 * c * s]
    x CAO
Conception assistée par ordinateur - Aspect social
Conception assistée par ordinateur - Logiciels
Corporation de réseau Computertime *voir* Réseau Computertime (Firme)
Coupe Canada (1987 : Tournoi de hockey) [27 * d * c]
    x Hockey - Série Canada-URSS
    x Matches de hockey Canada-URSS
    x Série Canada-URSS (Hockey)
    x Tournoi Canada-URSS (Hockey)
    x Tournoi de hockey Coupe Canada
Créativité dans les organisations [52 * l * s]
Critique dramatique [61 * c * s]
    x Critique théâtrale
Critique théâtrale *voir* Critique dramatique
Cytosine arabinoside [49 * l * s]
    x Ara-C

Up to recently, minimal human intervention (15 minutes or less for a 200-page publication) to supervise formatting done using Word macro command facilities

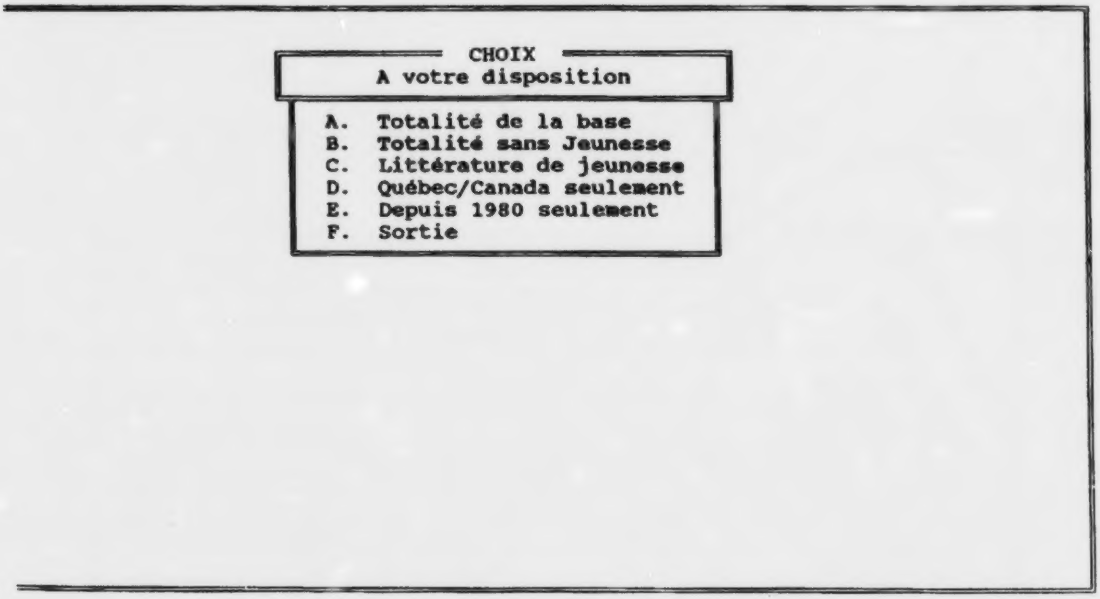
With Word 5.0 RFT (Rich-Text Format) codes directly embedded in PRINT files, with automatic conversion done by Word.

- Our CD-ROM experience

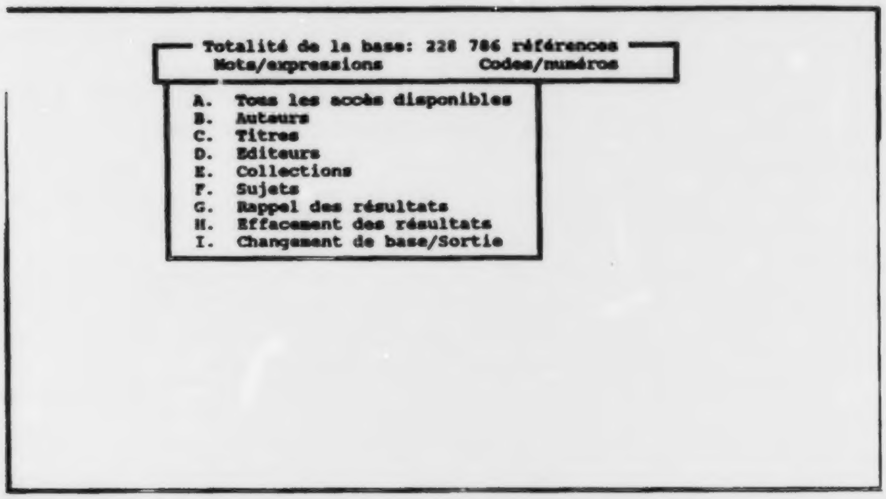
  Our first CD-ROM, including 228,000 records from our CHOIX database (books and serials) has been published in April 1989.



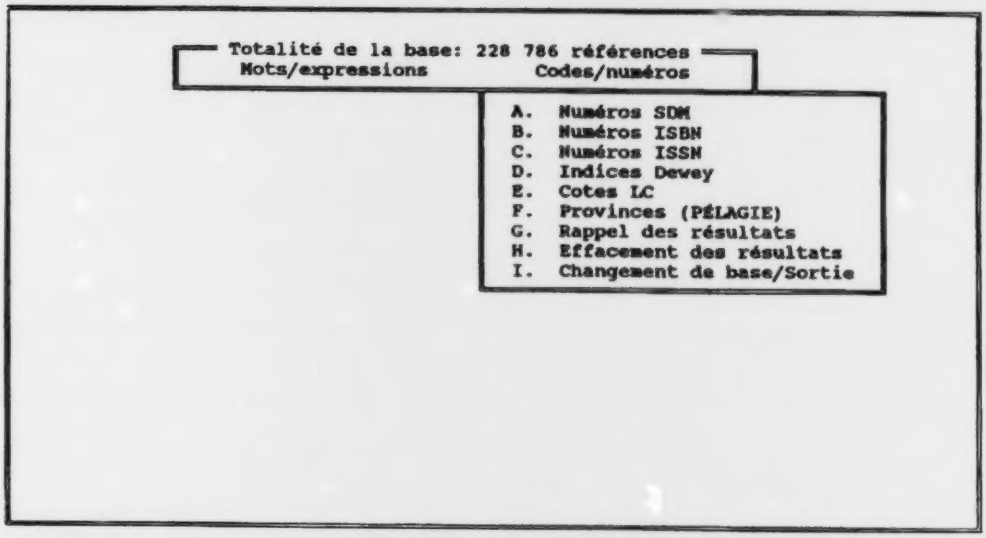
```
Services documentaires multimedia Inc.

CHOIX
documentation imprimée

Version CD-ROM
Janvier 1989

A votre disposition, 228 786 références
sur les livres et les publications en série
de langue française, dans tous les domaines

»»»»» Appuyez sur ↵ pour continuer «««««
```



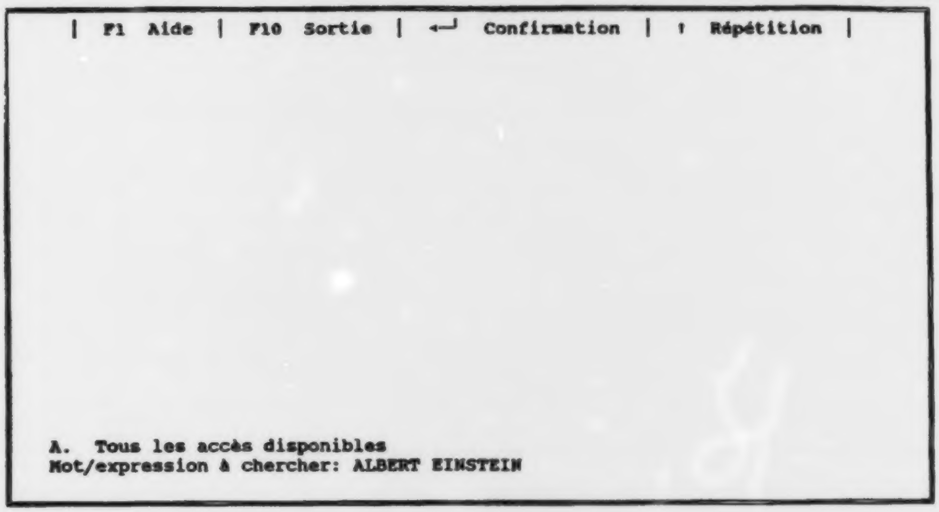
```
CHOIX
A votre disposition

A.  Totalité de la base
B.  Totalité sans Jeunesse
C.  Littérature de jeunesse
D.  Québec/Canada seulement
E.  Depuis 1980 seulement
F.  Sortie
```



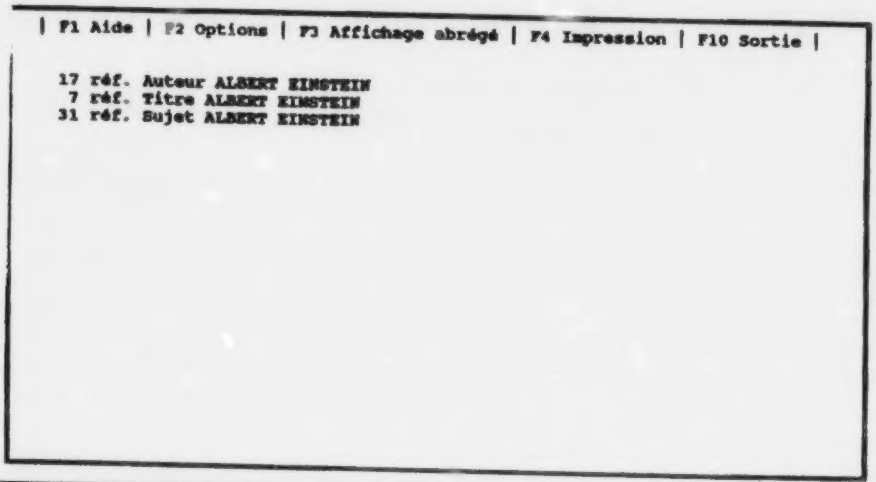
```
Totalité de la base: 228 786 références
  Mots/expressions          Codes/numéros

  A.  Tous les accès disponibles
  B.  Auteurs
  C.  Titres
  D.  Editeurs
  E.  Collections
  F.  Sujets
  G.  Rappel des résultats
  H.  Effacement des résultats
  I.  Changement de base/Sortie
```



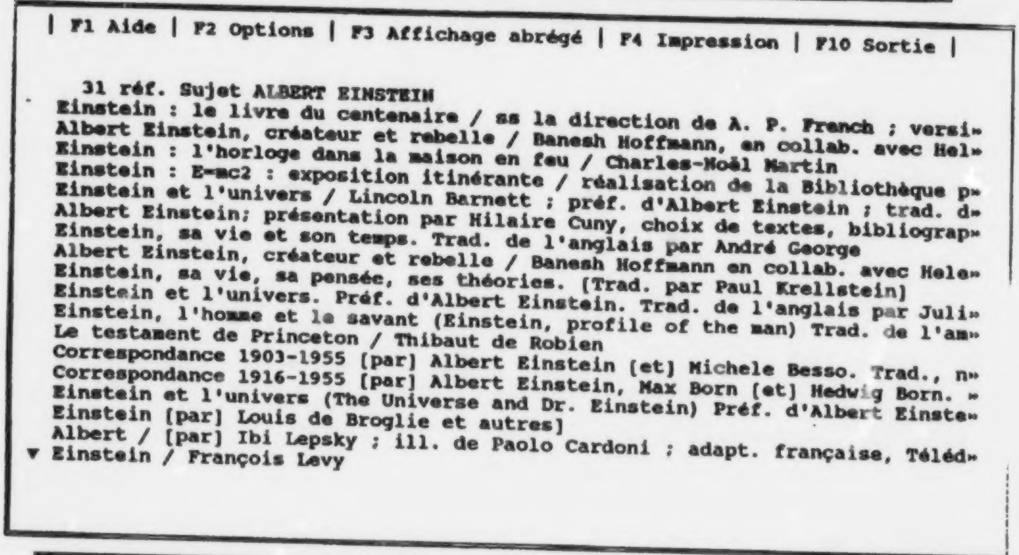
```
Totalité de la base: 228 786 références
  Mots/expressions          Codes/numéros

                          A.  Numéros SDM
                          B.  Numéros ISBN
                          C.  Numéros ISSN
                          D.  Indices Dewey
                          E.  Cotes LC
                          F.  Provinces (PÉLAGIE)
                          G.  Rappel des résultats
                          H.  Effacement des résultats
                          I.  Changement de base/Sortie
```



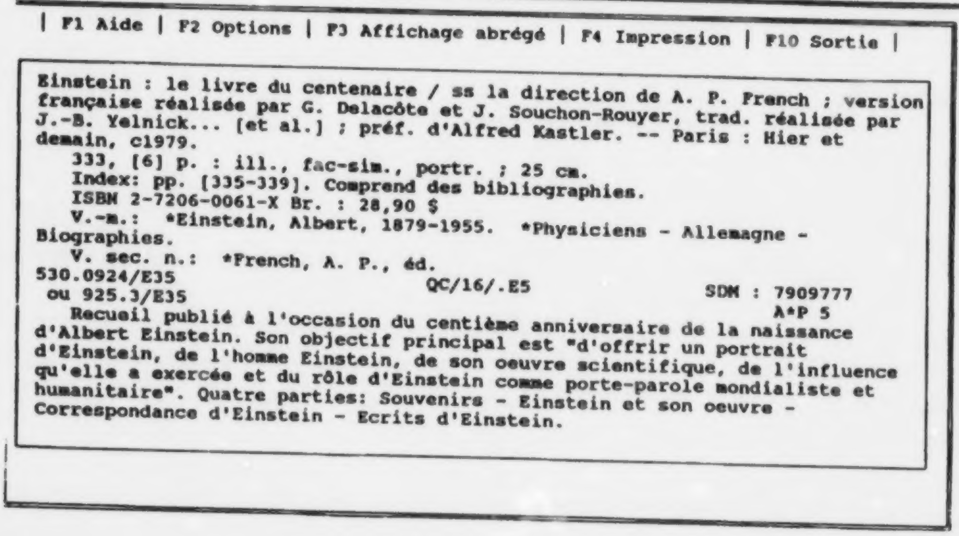
```
| F1 Aide | F10 Sortie | ←┘ Confirmation | ↑ Répétition |

A.  Tous les accès disponibles
Mot/expression à chercher: ALBERT EINSTEIN
```


| F1 Aide | F2 Options | F3 Affichage abrégé | F4 Impression | F10 Sortie |

```
  17 réf. Auteur ALBERT EINSTEIN
   7 réf. Titre ALBERT EINSTEIN
  31 réf. Sujet ALBERT EINSTEIN
```


| F1 Aide | F2 Options | F3 Affichage abrégé | F4 Impression | F10 Sortie |

```
   31 réf. Sujet ALBERT EINSTEIN
  Einstein : le livre du centenaire / ss la direction de A. P. French ; versi»
  Albert Einstein, créateur et rebelle / Banesh Hoffmann, en collab. avec Hel»
  Einstein : l'horloge dans la maison en feu / Charles-Noël Martin
  Einstein : E=mc2 : exposition itinérante / réalisation de la Bibliothèque p»
  Einstein et l'univers / Lincoln Barnett ; préf. d'Albert Einstein ; trad. d»
  Albert Einstein; présentation par Hilaire Cuny, choix de textes, bibliograp»
  Einstein, sa vie et son temps. Trad. de l'anglais par André George
  Albert Einstein, créateur et rebelle / Banesh Hoffmann en collab. avec Hele»
  Einstein, sa vie, sa pensée, ses théories. [Trad. par Paul Krellstein]
  Einstein et l'univers. Préf. d'Albert Einstein. Trad. de l'anglais par Juli»
  Einstein, l'homme et le savant (Einstein, profile of the man) Trad. de l'am»
  Le testament de Princeton / Thibaut de Robien
  Correspondance 1903-1955 [par] Albert Einstein [et] Michele Besso. Trad., n»
  Correspondance 1916-1955 [par] Albert Einstein, Max Born [et] Hedwig Born. »
  Einstein et l'univers (The Universe and Dr. Einstein) Préf. d'Albert Einste»
  Einstein [par] Louis de Broglie et autres]
  Albert / [par] Ibi Lepsky ; ill. de Paolo Cardoni ; adapt. française, Téléd»
▼ Einstein / François Levy
```


| F1 Aide | F2 Options | F3 Affichage abrégé | F4 Impression | F10 Sortie |

```
Einstein : le livre du centenaire / ss la direction de A. P. French ; version
française réalisée par G. Delacôte et J. Souchon-Rouyer, trad. réalisée par
J.-B. Yelnick... [et al.] ; préf. d'Alfred Kastler. -- Paris : Hier et
demain, c1979.
   333, [6] p. : ill., fac-sim., portr. ; 25 cm.
   Index: pp. [335-339]. Comprend des bibliographies.
   ISBN 2-7206-0061-X Br. : 28,90 $
   V.-m.:  *Einstein, Albert, 1879-1955.  *Physiciens - Allemagne -
Biographies.
   V. sec. n.:  *French, A. P., éd.
530.0924/E35
 ou 925.3/E35                QC/16/.E5                  SDM : 7909777
                                                             A*P 5
   Recueil publié à l'occasion du centième anniversaire de la naissance
d'Albert Einstein. Son objectif principal est "d'offrir un portrait
d'Einstein, de l'homme Einstein, de son oeuvre scientifique, de l'influence
qu'elle a exercée et du rôle d'Einstein comme porte-parole mondialiste et
humanitaire". Quatre parties: Souvenirs - Einstein et son oeuvre -
Correspondance d'Einstein - Ecrits d'Einstein.
```

We use a customized version of the SECONDE software package, developed by DESTIN Inc., a Montréal-based company.

Very simple user interface, using pull-down menus.

Full MARC output for libraries wishing to include records in their local database.

Batch module, for massive extractions of records, using our SDM numbers or ISBN numbers (up to 2500 records extracted per hour.)

For additional information:

Alain Boucher
Directeur de la Planification et du Développement
Services documentaires multimedia
1685, rue Fleury Est
Montréal (Québec)
Canada H2C 1T1

MINISIS AT PID

D S R MURTY & CHANDER SHEKHAR
Publication & Information Directorate

INTRODUCTION

Publications and Information Directorate (PID) is a constituent unit of the Council of Scientific and Industrial Research (CSIR), India. The two main functions of PID are (1) collection and dissemination of information on research and industry; and (2) Publication of scientific research journals. As a part of these functions, PID brings out (a) research periodicals in chemistry, physics, biochemistry and technology etc. (b) popular science journals (c) an encyclopedic publication called "The Wealth of India" (Raw Materials, and Industrial Products series), and (d) monographs on significant and socially relevant scientific topics.

A bimonthly publication, Medicinal and Aromatic Plants abstracts (MAPA) is being brought out by PID over a decade. It covers 400 important core journals relevant to the fields of medicinal and aromatic plants. Other related information services include: (i) Selective Dissemination of Information in the field of medicinal and aromatic plants (ii) Technical Enquiry Service (iii) Bibliography Service (iv) Statistics of Data, and (v) Document Copy Service.

COMPUTERIZATION AT PID

Some efforts have been made to computerize the information activities at PID since 1976. In this direction, a project has been undertaken to bring out the Keyword-out-of-Context (KWOC) indexes of the research periodicals of PID using successively the IBM-l620, IBM-360, Interdata 8/32, HP-1000 and HP-3000 systems from 1976 onwards. The software for this project was developed in-house using various lower and higher-level languages like SPS (on IBM-1620), FORTRAN IV, COBOL etc. on the IBM-360 and HP1000/3000 systems.

With the availability of a HP-1000 system at PID premises during 1961-83, some programs were developed for bringing out the keyword and author indexes of MAPA as well as to create a searchable Data base (DB) on medicinal and aromatic plants using the HP-1000 system. The programs developed were based on FORTRAN-IV and had some limitations regarding storage and retrieval methods owing to the limitations of the HP-1000 system which had only two tape drives with only a floppy drive and no hard disk drive. Thus the experimental DB was confined to storing the authors, keywords, titles, and affiliations, etc, without the abstract.

PID by virtue of its involvement in medicinal and aromatic plants information services has been identified by UNESCO as one of the potential information centres which can serve the purpose of a national node to collaborate in its Asian and Pacific Information Network in Medicinal and Aromatic Plants (APINMAP) project. In order to make a beginning in this direction, UNESCO, under its PGI programme, funded the purchase of a HP-3000/37 computer system. The system thus acquired has the following configuration:

CPU	512K
Cartridge Disk	55 MB
Cartridge Tape	55 MB
Line Printer	300 1pm
Laser Printer	8 p/m
Terminals	2
HP 150 PC	1

MAPA DATA BASE USING MINISIS

The MINISIS software package was installed at PID by Ms Mary Campbell of IDRC in August 1987. The training for the same was conducted during 10-21 August 1987 by Ms Campbell, and during 9-12 November 1987 by Mr. Bob McKercher

The MINISIS package was implemented during the end of 1987 for the creation of a DB on medicinal and aromatic plants introducing a number of fields to include a wide variety of bibliographic documents improving over the earlier experimental DB. The data definition table is given in Table 1. The hard copy is brought out for the mass production of MAPA through offset printing using a suitable print format which conforms to the earlier pattern when it was separately composed. The output for the hard copy is brought out by means of HP Laserjet Plus printer. Thus using MINISIS, the DB creation and the hard copy generation are integrated using one time key-boarding effort. The arrangement of the abstracts in the hard copy is according to a subject grouping of about 15 categories. The data are input through the UPDATE processor randomly as far as the category code is concerned. After the validation of data, the records are arranged in the sequence of this category code (4-digit) and within that according to the first author using INDEX processor. One field is allocated for the final abstract number in the DB. By running the RENUM processor, the final abstracts for one issue are generated for the hard copy printing. A sample page of the hard copy is given in Table 2 which shows the output generated by the print format picking up those field required for the hard copy of MAPA.

Using the INDEX processor, keyword and author indexes linking with the final abstract number are also generated and printed using suitable print formats.

HARDWARE CONSTRAINTS

Due to the limitations of the system configuration, mainly lack of significant hard disk space, at present it is not possible to keep the DB for online searching, inverting potential fields for retrieval. At present, no fields are inverted due to disk space limitation. This makes it necessary to take the data (batch files) on to cartridge tape(s) as the data grows.

SOME LIMITATIONS ENCOUNTERED IN MINISIS USAGE

The print format for the hard copy output is designed for two-column output with 38 character column width and 2 character inter-column gap. Since the print format breaks the line recognizing blank character only for this purpose, sometimes the lines generated leave much gaps to the right. The problem, it is felt, is much accentuated in the case of data dealing with scientific titles and abstracts where they contain long-winding chemical formulae sprinkled with a lot of hyphens and apostrophe characters running into 50 or 60 characters without a blank in between. Sometimes a line is generated containing only one word with two much blank spaces to the right. It is suggested that the print module may be modified to recognize some other characters (other than blank) like, comma, hyphen, colon, right parenthesis, one character before left parenthesis etc., so that the output looks much improved.

The use of KSAM file(s) (for validation of data even for one or two fields) has been found to slow up the data entry very much. It is hoped that future versions of MINISIS may do something to overcome this problem.

DATA EXCHANGE WITH APINMAP

As already mentioned, PID is acting as one of the National Nodes for the APINMAP project. The input from the different nodes in this network is being merged at the present Network Centre at AIBA using the micro CDS/ISIS package, the software adopted by APINMAP. At PID, the DB is created by using MINISIS package. From the DB thus created, the relevant records from INDIA and falling within the scope of APINMAP have to be sent to AIBA so that the same can be loaded to the APINMAP DB using CDS/ISIS. For this purpose, a correspondence definition is used to convert the data into ISO-2709 format for getting the output using ISOCONV processor given with the G version of MINISIS. It is particularly helpful where a provision has been made to get the output that is directly accepted by the CDS/ISIS in the form of 80 byte blocks adding the extra character for record separator. Using the QUERY processor, a hitfile generates the items of the Indian node from the total DB and the final output in the ISO format is obtained using the ISOCONV processor conversion.

OTHER PROJECTS USING MINISIS

Studies are being made to adopt MINISIS for the creation of some other data bases for use in PID. A DB is being envisaged for a comprehensive mailing list for the various journals of PID. Another project is underway for creating a DB containing the details of experts in various subjects for the internal use of editors.

Table 1 - Data Definition Table for the Medicinal and Aromatic Plants Database

*** Database : MAPA

FIELD NAME	MNEMONIC	TAG	OFFSET	LENGTH	NUMERIC	DEC PL	REPEAT	SUBFIELDED	PROMPT	MANDATORY
PARTICIPATING CENTRE RECORD NUMBER	PCRN	P100	-1	11					Y	
MAPA YEAR & NUMBER	NOYR	N010	-1	4					Y	
MAPA SUBJECT GROUPING CODE	SGC	S100	-1	4					Y	
GEOGRAPHICAL AREA CODE	GEOGAR	G100	-1	10			Y			
LANGUAGE CODE OF DOCUMENT	LANG	L100	-1	10			Y		Y	
ENTITY IDENTITY	DOCENT	D100	-1	45				Y	Y	
ENTITY CODE	DENTY	D101	-1	3					Y	
DOC TYPE	DTYPE	D102	-1	3					Y	
OTHER IDENTIFYING ELEMENTS	DOIE	D103	-1	3					Y	
IDENTIFICATION NUMBER	DIDNO	D104	-1	36					Y	
CONTRIBUTING INST CODE	CINSTN	C010	-1	20					Y	
PERSONAL AUTHOR(S)	NAME	N100	-1	60			Y		Y	
AFFILIATION	AFFLN	A010	-1	300					Y	
TITLE (ENGLISH)	TENGL	T100	-1	500					Y	
TITLE (OTHER LANGUAGE)	TOLANG	T200	-1	500						
SOURCE DOCUMENT	SDOC	H100	-1	500				Y	Y	
SOURCE PERSONAL/CORP AUTH	SPCATH	H101	-1	160						
SOURCE TITLE	STITLE	H102	-1	180					Y	
SOURCE EDITION	SEDN	H103	-1	20						
SOURCE IMPRINT	SIMPR	H104	-1	80						
SOURCE VOL ETC	SVOL	H105	-1	35					Y	
SOURCE YEAR	SYEAR	H106	-1	25					Y	
BIBLIOGRAPHIC NOTES ETC	REFETC	R100	-1	75					Y	
CORPORATE NAMES	CORPAU	C100	-1	250			Y	Y		
SUBORDINATE BODY	CSUBB	C101	-1	90						
INSTITUTION NAME	CINST	C102	-1	80						
COUNTRY	CCNTRY	C103	-1	80						
MEETING	MEETNG	M100	-1	300				Y	Y	
NAME OF MEETING	MNAME	M101	-1	160					Y	
MEETING PLACE	MPLACE	M102	-1	50					Y	
MEETING DATE	MDATE	M103	-1	30					Y	
MEETING ORGANISER	MORG	M104	-1	60					Y	
EDITION STATEMENT	EDNST	E100	-1	300						
IMPRINT	IMPRT	I100	-1	300				Y		
PLACE	IPLACE	I101	-1	50						
PUBLISHER'S DETAILS	IPUBLR	I102	-1	160						
IMPRINT DATE	IDATE	I103	-1	30						
COPYRIGHT	ICPYRT	I104	-1	60						
ANY NOTE/PRICE ETC	PRICE	P300	-1	150						
COLLATION	PAGES	P200	-1	150						
SERIES	SERIES	S200	-1	300				Y	Y	
TITLE OF SERIES	STITL	S201	-1	120						
STATEMENT OF RESPONSIBILITY	STRESP	S202	-1	70						
TITLE OF SERIES IN OTHER LANGS	SOTLNG	S203	-1	90						
FREQUENCY	SFREQ	S204	-1	20						
ABSTRACT	ABST	A100	-1	1500			Y		Y	
KEYWORDS	KEYWRD	K100	-1	100			Y		Y	
BOTANICAL NAME(S) OF PLANTS	BOTNAM	K200	-1	100			Y		Y	
NO FOR RENUM PROCESSOR	ABSTNO	A200	-1	4						
LOCATION OF DOCUMENT	LOCDOC	L200	-1	100						
PRIMARY SUBJ CATEGORY	PSBCAT	S300	-1	3					Y	
SECONDARY SUBJ CATEGORIES	SSBCAT	S310	-1	3			Y		Y	

Table 2 -A Sample page of the hard copy of MAPA

23, Mansoura, Egypt) Constituents of Conyza linifolia. Pharmazie, v. 42(5): p. 355, 1987 (4 ref, En)

beta-Farnesene, curcumene, lupeyl acetate, six acetylenes and germacrene D have been isolated from the aerial parts of the plant and identified.

8902-0774 Miyaichi, Y., Imoto, Y., Tomimori, T., Lin, C.-C. (School of Pharmacy, Hokuriku University, 3 HO, Kanagawa-machi, Kanazawa, 920-12, Japan) Studies on the constituents of Scutellaria species. 9. On the flavonoid constituents of the root of Scutellaria indica L.. Chemical and Pharmaceutical Bulletin, v. 35(9): p. 3720-3725, 1987 (17 ref, En)

Four new flavanones (I-IV) and a new flavone (V) were isolated from the root of S.indica, together with 2(S)-5, 2"-dihydroxy-7,8,6'-trimethoxyflavanone, (+-)-5,2'-dihydroxy-6,7,6'-trimethoxyflavanone, 5,7-dihydroxy-8,2'-dimethoxyflavone, rivularin, 5,2',6'-trihydroxy-7,8-dimethoxyflavone, scutevurin, 5,7,4'-trihydroxy-8-methoxyflavone, wogonin, alpinetin, caldamomin and wogonin 7-O-glucuronide. Compounds 1-V were identifed based on spectral data and simple chemical modifications. JICST, Tokyo.

8902-0775 Morales, V.F. (Department de Quimica, Facultad de Ingenieria, Universidad de Magallanes, Casilla 113-D, Punta Arenas, Chile) Chemical investigations of some Chilean Berberis species. Revista Latinoamericana de Quimica, v. 18(1): p. 46-50 , 1987 (17 ref, Es)

Benzyltetrahydroisoquinoline alkaloids with new natural patterns e.g. the new isoindolobenzazepine, isoindoloisoquinoline etc. and the new proaporphine and aporphine benzylisoquinoline and bisbenzylisoquinoline dimers have been described. Biogenetical aspects of these alkaloids and their relationship with other alkaloids of Berberis have been discussed.

8902-0776 Mossa, J.S., Cassady, J.M., Kozlowski, J.F., Zennie, T.M. , Antoun, M.D., Pellechia, M.G., MeKenzie, A.T. , Byrn, S.R. (Department of Pharmacognosy and Medicinal Plants Research Unit, Research Center, College of Pharmacy, King Saud University, Riyadh, Saudi Arabia) Novel 6,7-seco-6,11-cyclolabdane diterpenes from Cluytia richardiana. Tetrahedron Letters, v. 29(30): p. 3627-3630, 1988 (4 ref, En)

Two novel 6,7-seco-6,11-cyclolabdane diterpenoids, richardianidin-1 and 2, were isolated from the leaves of C. richardiana . The structure and stereochemistry of 1 and 2 were established to be 6-(acetyloxy)-15,16-epoxy-2,12-dihydroxy-6,7-seco-6,11-cyclolabda-8, 13 (16),14-triene-7,18-dioic acid gamma-lactone delta-lactone and 6-(acetyloxy)-15,16-epoxy-1,12-dihydroxy-6,7-seco-6,11-cyclolabda-8, 13 (16),14-triene-7,18-dioic acid di-delta-lactone based on spectroscopy and X-ray data.

8902-0777 Mukherjee, K.S., Chakravorty, C.K., Chatterjee, T.P. , Bhattacharjee, P. (Department of Chemistry, Visva-Bharati University, Santiniketan-731 235, WB, India) A triterpene from Salvia leucantha Cav.. Journal of the Indian Chemical Society, v. 65(6): p. 458, 1988 (6 ref, En)

Isolation and structure elucidation of a pentacyclic triterpene, 3 epi uvaol (mp 232 degrees) from the chloroform extract of S. leucantha has been described. Its structure has been demonstrated to be urs-12-en-3,28-diol.

8902-0778 Narayan, M.S., Bhattacharya, B., Dhanaraj, N. , Seshadri, R. (Department of Food Chemistry, Central Food Technological Research Institute, Mysore-570013, Karnataka, India) Free and bound triacontanol in tea leaves. Jounral of the Science of Food and Agriculture, v. 43(3), p.

Installation of Large Databases by Means of MINISIS

Wu Bing and Li Yanxia
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries

The name MINISIS is derived from MINI-ISIS. ISIS is a data base management system which uses large mainframe computers. MINISIS is a DBMS which operates on minicomputers. Now two large database have been built on minicomputer using MINISIS.

I. It is feasible to install large databases using MINISIS.

In 1986 and 1988, we installed two large document data bases, METADEX and COMPENDEX, separately on an HP 3000-68 computer. The two databases are the main source of foreign information on machinical and electrical industries. The METADEX data base passed the state test and appraisal organized by the Scientific and Technology Commission of China (STCC) at the end of 1986. It has been in test operation since then. The COMPENDEX data base will begin test runs at the end of this year.

The performance of the same data base produced on different DBMS can be compared. The information retrieval systems, such as DIALOG, ORBIT of U.S.A. and ESA-IRS of ESA, keep their own METADEX and COMPENDEX databases. On Dec. 27,1986, experts gathered by STCC evaluated the performance of the METADEX database installed on MINISIS: "It is at a practical level in terms of its capacity of documents, retrieval approaches, response time etc. The database has been developed to an advanced level among those retrospective search on-line retrieval systems using foreign tapes in China at present. The performance of the main retrieval approaches are almost the same as those in the same database of international on-line search systems."

Later, we improved the performance of METADEX database according to suggestions made by metallurgists. Certain functions on our METADEX database have surpassed the performance of the same database maintained by DIALOG. This point will be discussed in another paper. The main functions of the COMPENDEX database are the same for the versions (MINISIS and DIALOG) and the MINISIS version functions better in some areas than the ESA-IRS version. Through the efforts of the designers and researchers, this experiment has proved it is possible to build large databases based on MINISIS DBMS.

II. Brief Introduction to the Size of Two Large Databases.

1 METADEX database
METADEX is a document database originally built on an HP3000-68 computer using the F version of MINISIS. Each logical record occupies about 1000 bytes. There are about 500,000 logical records (1974-May 1987) in the database. The master file approaches the HP physical file limit of 532MB. The XREF file, various inverted files and auxiliary files occupy 155MB. The ratio of the inverted files, X-reference file and auxiliary files to the master files is about 0.3:1.

2 COMPENDEX database
COMPENDEX is a document database and works under MINISIS version G. Each logical record is about 1200 bytes. There are about 83,000 logical records (1982-1988) in the database. Because the version G provides a multi-master file function, the COMPENDEX master files are not limited by HP file size limits. At present, the COMPENDEX database keeps 4 master files, 3 of which are about 4096 X 80,000 bytes, 1 which is about 4096 X 19269 bytes. They occupy 1061MB of disc space.

III. The time span for building large databases.

Liberated from the hard work of manual searching, people save time using computerized retrievals. Building large document databases costs time. It is just a part of the work to load the source data into the computers. More time will be used building inverted files. The inversion of phrases or the whole field costs less than the inversion of single words. The inversion of phrases or the whole field in 20,000 logical records needs about 9 hours. Inversion using single words takes about 120 hours for the same 20,000 logical records. The inversion speed also depends on the host status. If CONDENSE maintenance is done the speed improves. We were surprized to find that the inversion speed for single words was halved after CONDENSE maintenance. The time spent on indexing single words is affected by the number of stop words selected. More stop words require more indexing time. Version G allows users to change KSAM type stop word files into B-tree files. When this is done, the indexing speed can be doubled under the same host conditions.

IV. File Management and Maintenance of Large Databases.

After a large database is built, it is deadly important to maintain the files of the database. If something is wrong in the file management the whole database may be damaged. Since the master file is very large, it is more easily damaged. METADEX database was built under the F version of MINISIS. The master file uses 532MB disc space. When the private volume discs do not work well, the master file is often affected. To compensate for this, a set of back-up tapes is prepared, including as many as 13 rolls of tapes. If any one roll among them does not work, the whole set of tapes means nothing. Although we have kept several copies of the master file, we can not guarantee that every roll is working well due to the instability of our HP magnetic tape drive. To solve this problem, we adopt disc to disc back-up. In version G, muli-master file function provides great convenience. In building the COMPENDEX database, we built 4 master files. The smaller master files are less prone to damage. If the back-up of one of the small master files did not work, it would be easy to change it.

V. Existing Problems.

During the installation of COMPENDEX based on MINISIS version G, the size of each master file is determined at the very beginning. An unresolved problem we've encounter is that when a logical record spans two master files, a read error message appears while this logical record in the database is accessed.

We have gained experiences in the practice of building large databases. MINISIS is a powerful tool in supporting large databases. We're very grateful to senior researcher Mr. Wu Kang and Mr. Zhang Xi for their advice on the project from the very beginning. Mr.

Richard Lee, Mr. Richard Palmer, Miss Mary Campbell and Ms. Ellen Ruygrok were a great help to us. Thanks go to all who helped us.

Reference:

MINISIS Manuals

DIALOG Information Retrieval Manual

ESA-IRS Information Retrieval Manual

HP 3000 MPE Manual

利用 MINISIS 建立大型数据库

吴　冰　　李艳霞

MINISIS 的原意是 MINI-ISIS．ISIS 是用于大型计算机的大型数据库管理软件．顾名思义 MINISIS 是用于小型机的小型数据库管理软件．但是，目前利用 MINISIS 已在小型机上建立了两个大型数据库．

一、用 MINISIS 建立大型数据库是可行的

我们于 1986 年和 1988 年分别在 HP3000-68 机上建立了两个大型文献型数据库《METADEX》（金属文摘）和《COMPENDEX》（工程索引）．这两个数据库是机电行业国外资料的重要来源．《金属文摘》数据库已于 1986 年底通过了由国家科委组织的鉴定．并已投入试用，《工程索引》数据库可望年底投入试用．

利用不同的数据库管理系统建立相同的数据库，对其功能是能够进行相互比较的．美国 DIALOG、ORBIT 情报检索系统及欧洲航天局的 ESA-IRS 情报检索系统都有《金属文摘》和《工程索引》数据库．1986 年 12 月 26 日，国家科委组织国内有关专家对用 MINISIS 建立的《金属文摘》数据库进行了鉴定，鉴定结果是:

"这一数据库从联机文献量、检索途径、响应速度等方面具有实用化水平，达到目前国内利用国外文献磁带建立的追溯联机检索系统的先进水平，主要检索途径也大致接近国际联机系统同一数据库的基本功能"．

尔后，我们又根据冶金专家提出的要求，对《金属文摘》数据库的功能进行了改善，使《金属文摘》数据库的个别功能超过了 DIALOG 情报检索系统中的《金属文摘》数据库的功能．这点将在另一篇论文中详细说明．用 MINISIS 建立的《工程索引》数据库的主要功能也接近了 DIALOG 系统中《工程索引》数据库的主要功能，而且个别功能超过了欧洲航天局 ESA-IRS 系统中《工程索引》数据库的功能．实践证明在数据库设计人员和 IDRC 工作人员的共同努力下，用 MINISIS 建立大型数据库是可行的．

二、两个大型数据库简介

1、《金属文摘》数据库

《金属文摘》数据库最初是在 HP3000-68 机上用 MINISIS F 版建立的．数据内容为文献型，每条逻辑记录约占 1000 字节，现已装入 1974-1987.5 月

文献约 50 万篇．主数据文件（master file）已接近 HP 机最大物理文件：532 兆字节，地址文件及各倒排文件和辅助文件占 155 兆字节，倒排文件、地址文件和辅助文件与主数据文件所占空间的比例是：0.3：1．

2、《工程索引》数据库

《工程索引》数据库也是文献型数据库，目前工作在 MINISIS G 版本下，每条逻辑记录约占 1200 字节，现已装入 1982–1988 年的文献 83 万篇．由于 MINISIS G 版本提供了多个主数据文件功能，使得《工程索引》数据库的建立不受 HP 机最大文件空间的限制．目前《工程索引》数据库已有 4 个主文件．其中 3 个主文件每个是 4096×80，000 字节，另一个主文件是 4096×19269 字节．共占磁盘空间 1061 兆字节．

三、建立大型数据库所耗费的时间

人们从繁重的手工检索方法中解放出来，利用计算机检索文献赢得了时间．但是在建立大型文献数据库时也要耗费大量时间．

在计算机里装入数据只是建库中的一部分工作．建立倒排文件要花费更多的时间．全字段倒排或词组倒排比以字为单位倒排花的时间相对来说少一点．全字段倒排或词组倒排 20，000 篇文献花费 9 小时，而以字为单位倒排 20，000 篇文献要花费 120 小时．倒排的速度还与主机工作状态有关．主机经常做 condense 维护，可提高倒排速度．使我们感到吃惊的是，有一次在主机做了收集碎片空间（condense）维护后，以字为单位的倒排速度竟提高了一倍．以字为单位抽词所花费的时间与所选禁用词的多少有关，禁用词多则抽词时间长．由于 MINISIS G 版本提供了将 KSAM 型禁用词库转为 B 树型的功能，经比较在同样数量的禁用词情况下（计算机条件相同），抽词速度提高了一倍．

四、大型数据库文件管理及维护

大型数据库建成后，数据库文件管理是至关重要的事．如果文件管理不当，花费大量时间建立起来的数据库就会毁于一旦．大型数据库的主文件很大，而文件越大，损坏的可能性也越大．《金属文摘》数据库是在 MINISIS F 版本下建成的，主文件占 532 兆字节．当私用卷磁盘机发生故障时经常会影响到这个文件．为此我们利用磁带做了一套转贮备份，一套共转贮 13 盘，其中

任何一盘有毛病这套备份就都不能用了．我们虽然有几套主文件备份，但由于我们的 HP 磁带机的稳定性不高，也不能保证每套磁带的每一盘都能正常工作．为了解决这个问题，我们采用磁盘芯转出备份．由于 MINISIS G 版本提供了多个主数据文件功能，在建立《工程索引》数据库时我们将主数据文件分为若干个小文件，文件小出错的机会少，若一个小文件的备份坏了，也容易替换．

五、存在的问题

用 MINISIS G 版本建《工程索引》数据库时，每个主文件的尺寸是事先规定好的．装库后，出现了一个逻辑记录跨跃两个主文件的情况．当从数据库中读这个跨跃两个主文件的逻辑记录时，便出现了主文件读错误的信息．这一问题尚待解决．

通过几年来建立大型数据库的实践，我们总结了一些经验．MINISIS 为建立大型数据库提供了有力的工具．建立这两个大型数据库得到了吴亢和张晰老师的指导．加拿大国际发展研究中心的李绰雄先生、李查德・铂玛先生、玛丽・坎堡和爱伦・拉格小姐都曾给过我们帮助，在此一并表示感谢．

Establishment of APT Database on MINISIS

Yu Xinsheng
Louyang Petro-Chemical Engineering Corporation

API 数据库在 MINISIS 上的建立

Establishment of APT Database ON MINISIS

洛阳石化工程公司　　余新生
Louyang Petro–Chemical Engineering Corporation
Yu Xinsheng

【内容提要】 本文分析了 API 磁带格式，介绍了 API 数据库的设计思想，以及如何将 API 磁带的数据装人 MINISIS 数据库。

【Abstract】 This article analyses API tape format, introduces the design of APT datase and the procedures for loading data from API tape into a MINISIS database.

一、API 磁带格式介绍

API 磁带是美国石油学会(American Petroleum Institute)提供的文献和专利磁带，从 1964 年开始。API 文献是从 150 多种期刊、40 余种会议论文集和其它出版物中摘编的，包括石油炼制和石油化学品、空气和水保护、储运、石油代用品、油田化学品等内容。API 专利则包括美、英和西德等国家的专利。

API 记录内容包括文摘号、文献题目、作者、出处和主题词等字段。API 磁带的格式不同于 SO2709 格式，磁带的密度为 1600BPI，代码为 EBCDIC，无标号，一个记录占一个物理块，长度不超过 6000 字节，每个记录各有一个头标区和数据区，没有目录区。

头标区长 18 个字节，该区有 7 个字段，它们的位置和意义如下:

字段位置	意 义
1	叙词个数的百位数
2～6	记录总长度的1／6
7～12	文摘号
13	标题字段的数目
14～15	作者字段的数目
16	文章出处字段的数目
17～18	叙词个数的十位数和个位数

数据区是由标题、作者、出处和叙词等字段组成的，每个字段都是由长度为 36 或 66 字节的块组成，不同记录中同类字段的出现次数不一样，故 API 记录是不定长的。数据区各字段的意义如下:

1.标题字段

每个记录至少有一个标题字段，每个标题字段的长度为 66 字节。

2.作者字段

每个记录最多可有 10 个作者字段，每个字段长 36 字节。

3.文献出处字段

一个记录最多可有 99 个出处字段，每个字段长 36 字节．

4.叙词字段

每个叙词占 36 个字节，一个记录最多可有 200 个叙词字段．叙词字段的第 1～33 字节为叙词本身，第 34、35 和 36 字节分别表示权符、职符和连符．

二、API 数据库的设计

API 数据库(APILIB)是利用 MINISIS 的 DATADICT 模块建立的．数据库是 RD 型的．DATADICT 的 DB 命令用于建立数据库的总体结构．首先定义数据库的头标信息，然后逐个定义各个字段属性．APILIB 一共有 5 个字段，其中叙词字段又是由 3 个子字段构成的．在检索 API 文献时，通常是通过叙词和权符去查询记录，很少通过作者和其它属性去查询，所以只需对叙词字段中叙词子字段(包括权符在内)进行倒排．

数据库的头标信息为:

DATA BASE NAME – APILIB
DATA MODEL NAME – APIDM
DATA BASE CLASS – RD
DATA BASE TYPE – 03
FILE TYPE – MASTER
GROUP NAME – PUB
DEFAULT QUERY FIELD TAG – F501
DEFAULT PRINT FORMAT FILE NAME – APIP

各个字段的定义如下:

FIELDNAME	MNEMONIC	TAG	LENGTH	REFEAT	SUBFLD
ABSTRACT NUMBER	ABSNO	F100	8	N	
TITLE	TI	F200	66	Y	
AUTHOR	AU	F300	36	Y	
SOURCE	SOUR	F400	66	Y	
DESCRIPTOR	DSCR	F500	36	Y	Y
DESCRW	DESW	F501	34		
ROLE	R	F503	1		
LINK	L	F504	1		

三、API 磁带的预处理

因为 API 磁带不同于 ISO2709 格式，所以 MINISIS 不能直接接受 API 记录，而需要进行预处理．MINISIS 没有提供这类预处理软件，故这部分工作要由我们完成．由于 MINISIS 可以以 BATCHIN 方式接受数据，所以我们可以设计一个程序，把 API 磁带上

的记录转换成 BATCHIN 格式的文件．这种方法较之把 API 数据转换成 ISO2709 格式，要来得简便．

对于 APILIB 数据库，BATCHIN 模块能够接受的 API 记录格式应该是这样的：

```
DB=APILIB
!CHANGE STRIP=Y,LENGTH=72,LIST=Y
!REC ID
!F100!31F003
!F200!EFFECTIVE COMPONENTS OF DRILLING MUDS
!F300!KONOVALOV E A
!F300!SPIRIDONOV O N
!F400!GAZOV,PROM-ST,N,8 26-7  %1983<  IN$   RUSSIAN
!F400!*  *LENANEFTEGAZGEOLOGIYA,USSR
!F500!F501!-GEMENTING MATERIALS                    D
!F500!F501!-FORMATION CONTROL AGENTS               D
!F500!F501!-GENERAL ASPECTS,DRILLING FLUIDS        D
!F500!F501!ACCELERATOR                             N!F504!B
!F500!F501!ADDITIVE                                D!F504!C
!F500!F501!ALUMINUM SULFATE                        D!F504!C
!F500!F501!ASBESTOS                                D!F504!D
!F500!F501!CEMENTING                               D
!F500!F501!DRILLING FLUID                          *
!F500!F501!PLUGGING AGENT                          D
!F500!F501!SODIUM SILICATE                         D!F504!B
!F500!F501!THICKENER                               P!F504!D
!F500!F501!WELL PLUGGING                           P
```

转换程序的流程是：

1.打开 BATCHIN 格式输出文件，写入文件头信息：

```
!DB=APILIB
!CHANGE STRIP=Y,LENGTH=72
```

2.把记录头标志"!REC-IC"写入该文件；

3.从磁带上读入一个记录，分离出各个字段，加工成要求的格式，写到该文件去；

4.若 API 磁带未处理完，则转第 2 步，否则结束．

四、API 数据的装入

预处理后的 API 记录，形成了 BATCHIN 格式的文件，该文件可以是磁带文件或磁盘文件，这时，可输入命令：

```
:FILE   BATCH;    DEV=TAPE(或 DISC)
```

进入 MINISIS 的 BATCHIN 模块，将 BATCH 文件传至 APILIB 数据库中．上述过程可用一个作业来描述:

```
!JOB APIBATCH，MR.LM
!FILE BATCH;DEV = TAPE
!MINISIS
BATCHIN
APILIB
EXIT
!EOJ
```

所有的 API 记录都进入 APILIB 之后，再把 F501 子字段从中抽出，排序后进行脱机倒排，产生出叙词子字段的倒排文件．

FISH GERMPLASM RESOURCES DATABASE APPLICATION SYSTEM (FGRAS)

Ni Rufeng, Huang Nenggeng
Freshwater Fisheries Research Centre,
Chinese Academy of Fisheries Sciences

ABSTRACT

Freshwater fish germplasm resources database application system can guarantee the correctness of related data through self designed thesaurus of fish species name and the inverted file of ASFA, which serves as authority file. By application of interface technique of V/3OOO and ADVANCELINK, it solves the problem of data shared by different systems. It also improves mathematic statistics functions of the system by means of self defined application procedure. This paper describes the configuration of FGRAS and the main functions and performance characteristics.

BACKGROUND

There is a wide variety of freshwater fish species in China. It holds the wealth of fish germplasm resources. As the research on freshwater fisheries and aquaculture develops, data on fish physiology, ecology and cytogenetics increase and urgently need scientific management by use of computer as a tool on the basis of genetics theory. We set out to design the system in 1986. In this system, we compile the data of morphology, physiology, ecology, molecule genetics, immuno-genetics and biochemical genetics of the main economic fishes in China and input them into computer for sorting and statistical analysis.

DATABASE SYSTEM CONFIGURATION

The fundamental composition of FGRAS are shown in Fig 1 and 2. The logic structure of the database is shown in Fig 3. It consists of five basic relational databases (RD's) and one graphic and plate file system. It has 550 data fields among which the physiological, ecological and cytogenetic database are composed of several projected subset databases (PS's).

RD 1 physiological and ecological DB (PHYECO)

(1)	PS. E1	Morphological data (MORPHO)
(2)	PS. E2	Age and growth data (AGEGRO)
(3)	PS. E3	Sexual maturity age and reproductivity (SEXRE)
(4)	PS. E4	Data on adaptability to environment (ADAPEN)
(5)	PS. E5	Muscle composition data (MUSCOM)
(6)	PS. E6	Digital genetic analysis data (DIGGEN)

RD 2 Cytogenetic DB (CYTGEN) master

(1)	PS. C1	Karyotype analysis data (KTA)
(2)	PS. C2	DNA and t-RNA analysis data (DNA, t-RNA)
(3)	PS. C3	Electrophoretic analysis data (ELECA)
(4)	PS. C4	Isoenzyme analysis data (ISOENZ)
(5)	PS. C5	Gene frequency analysis (GENFRE)
(6)	PS. C6	Genetic analysis data (GENAN)

RD 3 Morphological ontogeny DB (master)

RD 4 Genetype ontogeny DB (master)

RD 5 Thesaurus of scientific names of fishes DB

RD 6 Graphic and Plate file system

THE MAIN CHARACTERISTICS AND FUNCTIONS OF THE SYSTEM

1. FGRAS is more practicable and it not only contains the original functions of MINISIS and V/3OOO, e.g. all the processing support recording up to 64000 characters, use the thesaurus of data, maintain the definition of data by database management, provide defined inter-account DB index, optimize data storage and retrieval operation, support full screen form input and repeat the field up to 800 times, etc., but also has mathematic statistical functions, e.g. by use of gene frequency analysis and other related data an automatically calculate the genetic distances, genetic correlation coefficient and genetic ability between different species to provide a theoretic basis of rational disposition for geneticists and policy-makers to choose an optimized design.
2. In application of JOIN and PROJECT operations of MINISIS relational database, not only multi-database joint retrieval can be performed but various user views will also be set up on basic RD for users and the controlled condition to DB access and fast access will be defined for users. This saves time for user to access as well as guarantees the integrity and security of database.
3. To build the thesaurus of scientific names, common names and local names of fish and to define its broader terms, narrower terms and related terms will facilitate the location of various fishes in taxonomy and the discrimination of their genetic relationship and provide translation from 3 language to another in Latin, English and Chinese.
4. Because of multidisciplinary involvement in germplasm resources information and difficulties to read and to memorize many Latin data fields, the interface technique of MINISIS and V/3OOO was adopted to make full screen editor so that data input formats will comply with user data sheets and provide Chinese suggest. Common key fields are easily and visually input by means of code selection mode.
5. The related inverted file of ASFA and AQUIS which serve as authority files of FGRAS will guarantee the correctness of key fields.
6. A variety of output formats: the output may be printed out or recorded on 550 2709 magnetic tape or dBase-III magnetic disc according to users' request. This makes it a common use of database for users at home and abroad.

THE FUNCTIONS OF THE SYSTEM

FGRAS adopts three levels of security measures so that the data are sound and reliable.

User	DBMS	File
User standard evaluation by use of password	Access control for user ability discrimination	Crypto access

The correctness and compatibility of data are realized through the related characteristics of MINISIS and V/3OOO, e.g. to link-define data access; to define the restricted condition of PS, DS; to build authority file and to define the type, scope, accuracy and specification of data acquisition.

Moreover, FGRAS is still put under the environment of DBMS of MINISIS. Once the system is destroyed by malfunction, the system can be restored to the normal state by use of system log file, periodic SYSDUMP backup magnetic tape.

In summary, there is much room to be improved in FGRAS, e.g. data amount, languages and fish species of thesaurus file could be enlarged in future; optimized breeding dynamic analog can be added to application procedure. We sincerely hope IDRC and all the colleagues at home and abroad will, as always, give us technical help and support.

FIG.1

FGRAS CONFIGURATION

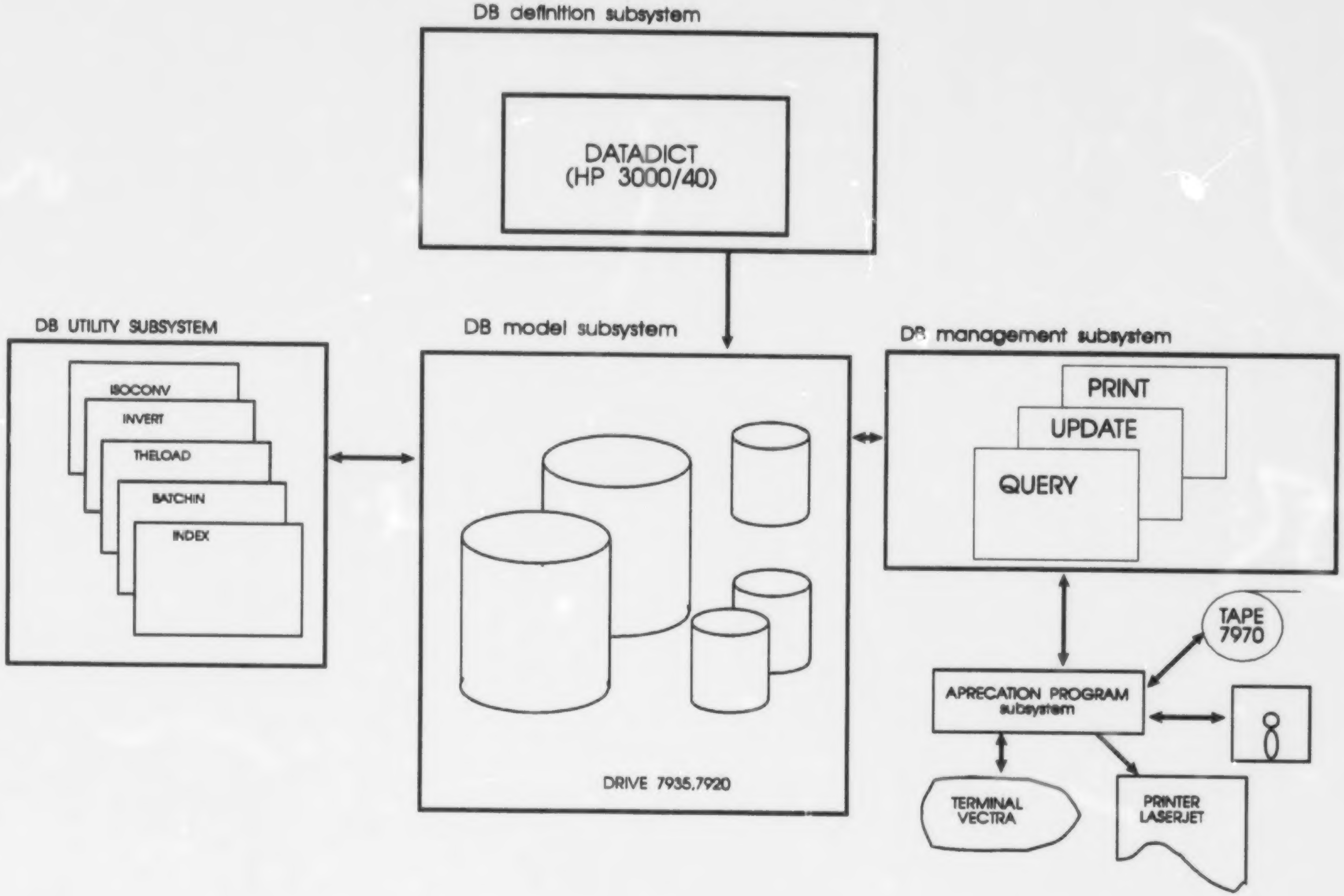

DB definition subsystem
DATADICT
(HP 3000/40)
DB UTILITY SUBSYSTEM
ISOCONV
INVERT
THELOAD
BATCHIN
INDEX
DB model subsystem
DRIVE 7935,7920
DB management subsystem
PRINT
UPDATE
QUERY
TAPE 7970
APRECATION PROGRAM subsystem
TERMINAL VECTRA
PRINTER LASERJET

Fig. 2 FGRAS FUNCTION STRUCTURE

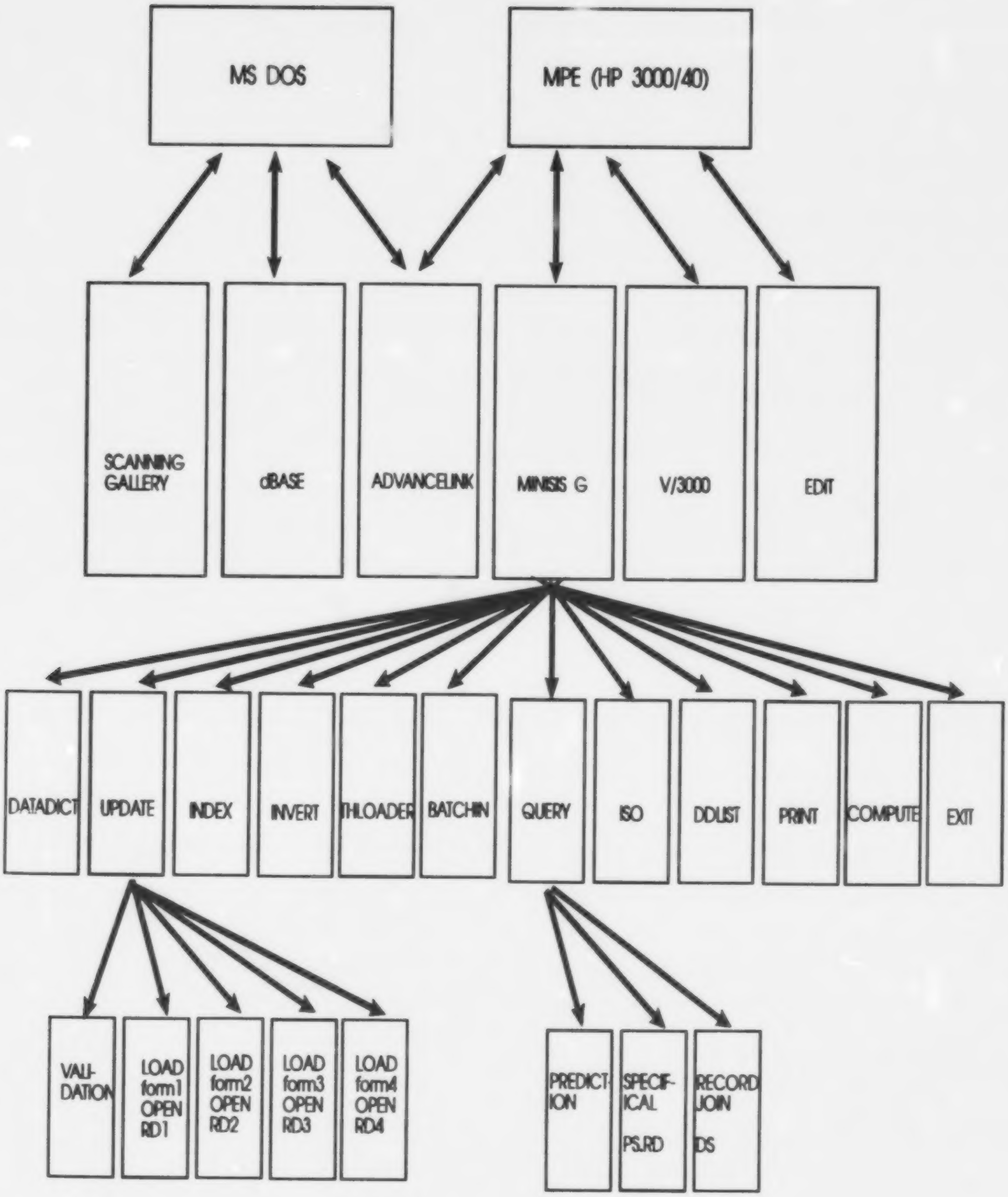
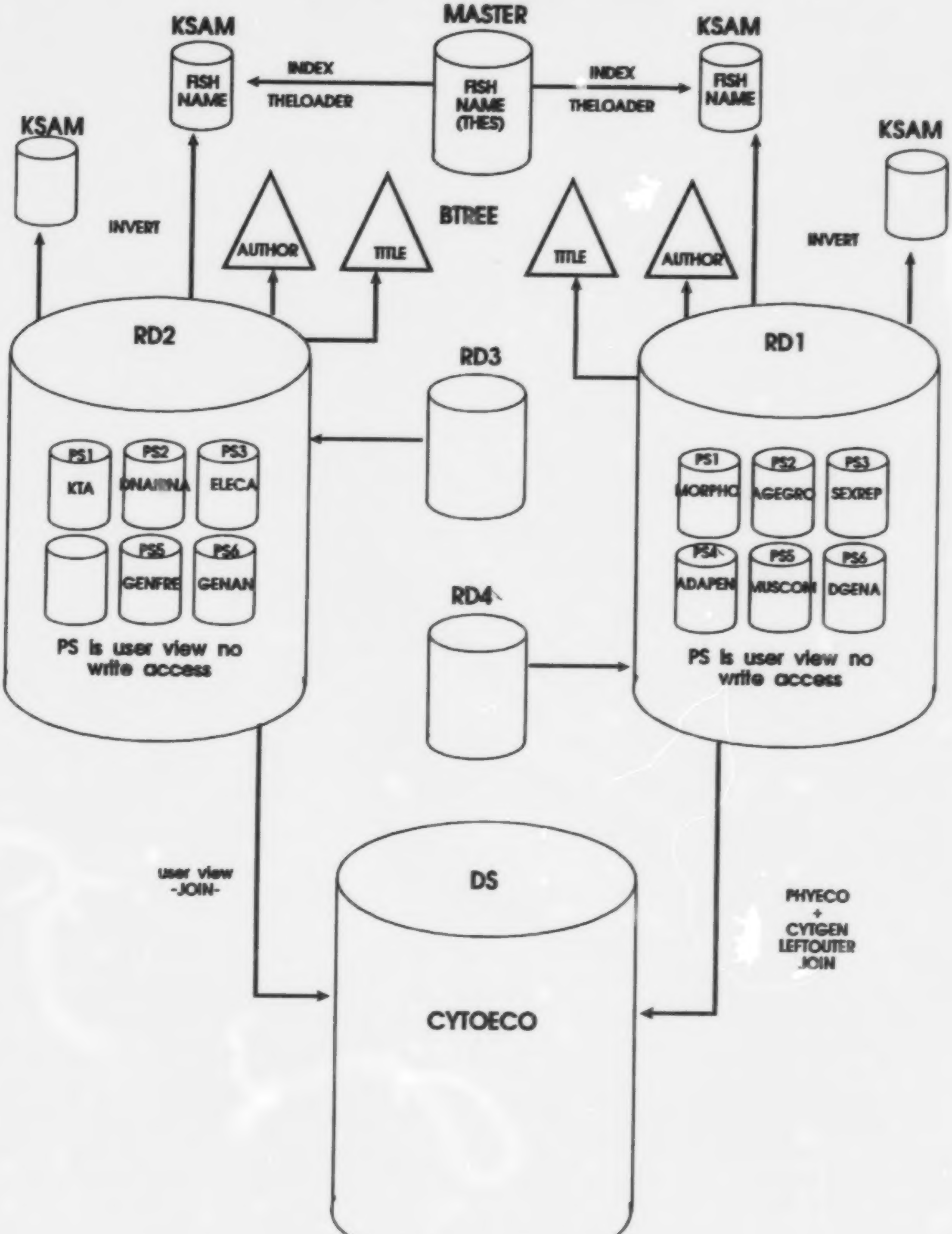

MS DOS
MPE (HP 3000/40)
SCANNING GALLERY
dBASE
ADVANCELINK
MINISIS G
V/3000
EDIT
DATADICT
UPDATE
INDEX
INVERT
THLOADER
BATCHIN
QUERY
ISO
DDLIST
PRINT
COMPUTE
EXIT
VALI-DATION
LOAD form1 OPEN RD1
LOAD form2 OPEN RD2
LOAD form3 OPEN RD3
LOAD form4 OPEN RD4
PREDICT-ION
SPECIF-ICAL PS.RD
RECORD JOIN DS

FIG. 3

FGRAS LOGICAL DB STRUCTURE


KSAM
MASTER
KSAM
FISH NAME
INDEX
THELOADER
FISH NAME (THES)
INDEX
THELOADER
FISH NAME
KSAM
KSAM
INVERT
AUTHOR
TITLE
BTREE
TITLE
AUTHOR
INVERT
RD2
RD3
RD1
PS1
KTA
PS2
DNA/RNA
PS3
ELECA
PS5
GENFRE
PS6
GENAN
PS is user view no write access
RD4
PS1
MORPHO
PS2
AGEGRO
PS3
SEXREP
PS4
ADAPEN
PS5
MUSCON
PS6
DGENA
PS is user view no write access
user view
-JOIN-
DS
CYTOECO
PHYECO
+
CYTGEN
LEFTOUTER
JOIN

鱼类种质资源数据库应用系统
Fish Gemplasm Resouces Data Bace Aplication System (FGRAS)

（中国水产科学研究院淡水渔业研究中心　倪汝凤　黄能耿）

摘要（ABSTRACT）

淡水鱼类种质资源数据库应用系统，通过自行设计鱼类种名的词表文件及引用ASFA的倒排文件作为权威文件（authority file）。保证了有关数据的正确性。通过应用V／300的接口技术和ADVANCELINK功能解决了不同系统的数据共享问题。通过自编的应用程序，增强了系统的数理统计功能。本文介绍了FGRAS的构成及主要的功能和性能设计特点。

一、引言（INTRUDUCTION）

我国淡水鱼类品种繁多，种质资源十分丰富，随着水产研究的不断深入，鱼类生理生态学及细胞遗传学方面的数据日益增多，迫切需要以计算机为工具，运用遗传学理论对这些数据进行科学的管理。故于１９８６年开始进行系统调研和分析。系统重点收集、加工我国主要经济鱼类的形态学、生理生态学、分子遗传学、免疫遗传学和生化遗传学等方面的数据输入计算机，用遗传学理论进行分类和统计分析。

二、　系统的组成（DBSYSTEM COMFIGRATION）

FGRAS基本组成见图一、二，库中数据的逻辑结构见图三。它有5个基本RD和一个图形文件集组成，共併包含５５０个数据项（fields）。其中生理生态库和细胞遗传库又分设若干个PS。

1、 RD．1 生理生态数据库——PHYSIOLOGICAL & ECOLOGICAL DB (PHYECO) (MASTER)

  (1) PS．E1——MORPHLOGICAL data (MORPHO)
  (2) PS．E2——Age & Growth data (AGEGRO)
  (3) PS．E3——Sexnal Maturity Age and Reproductivity data (SEXREP)
  (4) PS．E4——Adaptability to Enviroment data (ADAPEN)
  (5) PS．E5——Muscle Composition data (MVSCOM)

2、 RD．2 细胞遗传数据库 CYTOGENETICAL DB (CYTGEN) MASTER

(1) PS. C1——Karyotype Analysis data (KTA)
(2) PS. C2——DMA & tRNA Analysis data (DNATRNA)
(3) PS. C3——Electrophoretic Analysis data (ELEC)
(4) PS. C4——Isoenzyone Analysis data (ISOENZ)
(5) PS. C5——Gene Freguency Analysis data (GENFRE)
(6) PS. C6——Genetical Analysis data (QENAN)

3、 RD. 3 形态学个体性状数据库 (MASTER)
4、 RD. 4 基因形个体性状数据库 (MASTER)
5、 RD. 5 鱼类名称词表数据库 (THESAURUS)
6、 图谱文件集 Graph and photo file system

三、 系统特点和主要功能（SYSTEM CHARACTERS AND MAIN FUNCTIONS）

1、FGRAS实用性强，不仅保留了MINISIS和V／300的原不功能。如：所有处理程序支持长达64000个字符的记录，使用数据词典，以库管理方式维护数据定义，提供跨帐号库定义检索和数据存取优化，支持全屏幕数据输入，重复字段可达800次等。还具有较强的数理统计功能。如：利用基因频率分析和其他有关数据自动计算不同种之间的遗传距离、遗传相关系数和遗传力等为遗传工作者和有关决策部门提供优育种方案选择和品种合理布局的理论依据。

2、应用MINISIS关系型数据库的并操作和投影操作，既可实现多库联结检索，又可对基本RD设置各类用户索窗以定义库访问的约束条件和用户快速访问属干性。这不仅节省了用户访问时间又保证了库的完整和安全。

3、建立了鱼类名称的拉、英、汉词　表文件　、定义了其上位词、下位词，为确定各种鱼类的生物学分类位置及其关亲缘关系提供了方便；为鱼类名称提供了拉、英、汉对译。

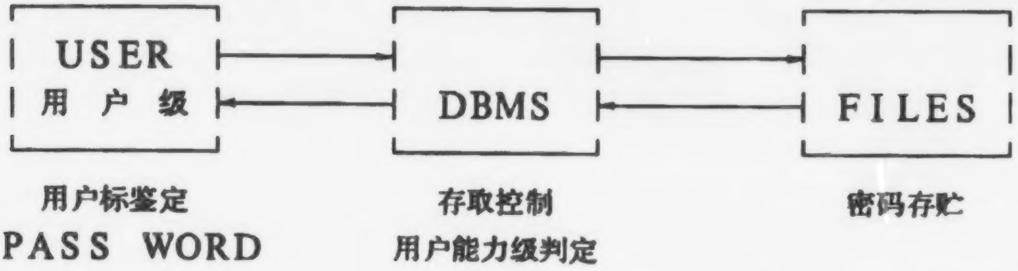
4、由于种质资源信息涉及专业面广。且许多数据项名采用拉丁文缩写难读难记，采用MINISIS与V／300接口技术，进行全屏幕编辑，使录入数据格式与用户数据调查表一致，并提供了中文提示。对常用关键字段采用代码选择方式录入作到了简便直观。

5、由于借用了ASFA及AQUIS的有关倒排文件作为FGRAS的权威文件，保证了关键字段什的正确可靠性。

6、输出格式多样化。按用户要求可以是表格式打印输出，也可以是ISO2709磁带和 dBASE 磁盘方式输出。使数据能与国内外用户充分共享。

四、系统的性能（SYSTEM SPECIFICATIONS）

FGRAS采用三级保密措施，以确保数据的完全可靠。



数据的正确性和相容性，则是利用V／300和MINISIS的有关特性，如：通过对中访问的链定义，对PS、DS的约束文件定义、建立权威文件以及对数据取什类型、范围、精度和SPECIFICATION等作严格的规定来实现的。

另外．由于FGRAS仍处于MINISIS的DBMS环境下，一旦系统遭到破坏,可以利用系统日子文件和定期转（SUSDUMP）的后缓磁带重新入库（RESTORE），即可使系统恢复到故障前某一时刻的正常状态。

结束语

FGRAS目前还有许多不中之处。如：数据量、词表文件的语种及品种数据量都还有待于不断扩充；应用程序还需增加优育肿动态模拟等功能。我们希望IDRC和国内外同仁能一如继往地在技术上给予帮助和支援。

Establishment of Japanese Document Retrieval System Based on MINISIS DBMS

Wu Kang
Zhou Jiang
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries

Abstract

This paper discusses our research into how to retrieve Japanese documents conveniently into a Chinese database built using MINISIS.

Japan's < Current Bibliography on Science and Technology >, CBST, covers international scientific literature. Many scientific research organizations and libraries collect the books listed in the bibliography. Although there is no denying that computerized information retrieval is superior to hand-retrieval, no other organization in China has CBST. This is due to the difficulty of processing Japanese characters in existing Chinese systems. The Japanese character set has different requirements than the Chinese systems. There also is a language barrier for Chinese users performing interactive searches.

MINISIS provides rather flexible interface conditions for building a Japanese literature data base. This flexibility allows us to improve the man-machine relationship.

In 1984, we bought the CBST tape in English-katakana mode. The English-katakana mode uses the EBCDIC character set. The highest bit of katakana in the set is "1." Using those tapes, we began to build a database using MINISIS DBMS, on HP-3000 series computers. HP line printers produced the katakana characters at high speeds. The data is convenient to process and fast, but replacing the square characters with their phonetic letters posed difficulties for Chinese technical men.

We introduced the CBST tape in kanji-kana mode to improve the man-machine relationship. We have studied the kanji-kana characteristics continuously and input the data into the database. Now there are about 50,000 Japanese scientific literature titles in the database with nine inverted fields. The database occupies about 400 MB in HP hard disc, with no problem on displaying and printing.

Data is input using the "kulo moto" method in Japan. This allows katakana or kanji to be input directly from the keyboard. The display is katakana or kanji during interactive searching. This is easy for the Japanese because they pronounce their home language well. It is quite embarrassing for our technicians to try to pronounce Japanese. They can read and understand kanji rather accurately if they study a little Japanese grammar. The Japanese kanjis are imitations of the Chinese square characters, including the stroke and the meaning. For these characters, only the pronunciation is different. The Chinese phonetic letters, "Pinyin," can replace the Japanese katakana for searching kanji. This eliminates the language barrier for Chinese users.

What problems would we encounter if we did so? What should we use to identify and replace the kanji developed in Japan? How do we deal with the problems of grammar?

To begin to answer these questions we did extensive mathematical statistical investigations on the internal character set of Japan International Centre of Science and Technology. We selected this set because it would contain all kanjis that would appear in CBST. The study showed that the kanjis with the same stroke and the same meaning as Chinese square characters make up the most of the set. Kanji with slight ratio in stroke but the same meaning with Chinese square characters make up a small ratio in the set. Kanji with the same stroke but different meaning and the kanji with both different stroke and meaning appear very, very infrequently. The statistics of frequency of the occurrence of each kind of kanjis in scientific literatures in Japanese changes in the same tendency. Simply, the bigger the difference the rarer the word. The results are summarized in table 1.

Table 1. The characteristics of kanjis in scientific literatures in Japanese

Category No.	Stroke	Meaning	Ratio in set	Frequency
1	Yes	Yes	938	970
2	No	Yes	58.7	20
3	Yes	No	0.2	3
4	No	No	3.1	7

Note: Yes – meaning those kanjis being the same as Chinese.

No – meaning those kanjis being different from Chinese.

Based on the above statistics, we can see that the kanjis in the first and the second category are the majority in their character set. The frequency of their application is also the highest; there will be no problem if we replace kanjis with Chinese in these two categories because the words are almost the same stroke and the same meaning. This not only saves us time and labour, but also increases the confidence in "Pinyin" searches. Using an exit from the query processor, we can translate the characters in the third character. In the exit, a check list matches one Chinese character or phrase to a Japanese character or phrase for the purpose of translation. Finally, we only adopt Japanese phonetic writings, the katakana for those few kanjis in the fourth category. They are few and their pronunciations are easier to remember or to find in a list. Phrases are treated the same as the kanjis in the third category. The amount of work is greater for each succeeding category. If the input is simply the phrases that correspond with Japanese custom and culture, the query time will be shorter.

After the above treatment, the Chinese users then can retrieve Japanese Scientific literature from within the framework of Chinese custom and culture. Because the feasibility and the flexibility of MINISIS DBMS our man (here Chinese)-machine relationship is improved.

To reduce input operations during query processes, some special file preparations are done to the BATCHIN programs. The basic principals are:

1. Replace all the words in Japanese character set with ASCII characters if there is a correspondence for those inverted fields.
2. The data on tape are based on 2 character sets (EBCDIC HITACHI 8000 series, JICST internal Japanese character set). By running the

program we changes them into ASCII code or the internal Chinese character code of our machine. We also process the labels for showing the combined words, boldface type and italics to suit the HP I/O system.

To speed the process up, data in main memory keeping, bit operation, recursion and data blocking are adopted in the program. Some separations or combinations for some phrases or fields are also prepared for later convenience in processing.

It is the heart of this research that interactive query on MINISIS use Chinese "Pinyin" instead of Japanese katakana to search the scientific literature written in Japanese.

During the research, Dr. Anthony Debonse, the dean of the department of Information Science and Library Science of Pittsburgh University, U.S.A., has advised our project.

References:

1. MINISIS manuals

2. HP 3000 computer language & subsystems

3. HP Chinese I/O systems

4. Information retrieval systems - characteristics, testing and evaluation F. Wilfrid Langcaster

5. Fundamentals of Chinese language computing

6. JIOS active manuals.

《用 MINISIS 建立日文文献检索系统》

吴亢　　　　周江

文摘： 作者重点指出了一条利用中文方式通过 MINISIS 数据库管理系统方便地检索日文文献的新途径．

日本《科学技术文献速报》是世界上著名的大型综合检索工具，我国许多科研机构和图书馆都收藏着它的书本式分册．但是国内其它单位迄今还没有建立它的日文汉字计算机检索系统．这主要因为日文文字在现有的中文系统上开发困难较大；一般的大型数据库管理系统的调节功能不够灵活；需要增加专门的日文汉字 I／O 设备等等．

MINISIS 为建立日文文献数据库提供了灵活的接口条件．借此手段，能够较为理想地改善人–机对话的环境．

1984 年起，我们引进了《科学技术文献速报》英文、片假名版磁带．此带采用的是 EBCDIC 代码表，表中的片假名是通过高位置“1”的方法表示的．以此为标志，我们开始了利用 MINISIS 在 HP–3000 系列计算机上建立日文文献检索系统的工作．能在 HP2563A 行打机上高速打印日文片假名文献．

这种数据的优点是处理方便，快捷，但其缺点是以音代字，不直观，特别是对中国技术人员来说，理解困难较大．于是我们自 86 年起开始引进日文汉字版的《科技文献速报》，开始了对日本汉字特点的研究工作并继续用 MINISIS 建立日本文献的数据库．现在，数据库中装有约 5 万篇的日本科技文献，能按九个倒排档检索，占用磁盘空间 400MB 左右．显示打印日文汉字无问题．

日本人检索，采用“黑本”输人法，在键盘录人表示日语语汇发音的假名，联机看到的也是假名或日文汉字．拼写和理解假名，对普通中国科技人员说来极其困难．这里所说的“普通”，不是指对工业技术的不精通，而是指对日文发音和将其作为准确的检索手段不精通．事实上，我国大部分工业技术人员都是这样．如果对检索词的绝大部分都需要查其假名发音，这对我们则是一个沉重的负担．而另一方面，尽管如此，凡学习过一点日语的人都能较准确的阅读和理解日文，这是因为日本汉字最初是从中国传人的结果．许多日本文字的笔划、意义都是照搬中国的，只是其发音较我国普通话差别很大．

如果能用汉语拼音代替假名作为一种检索语文来查询，就为我国科技人员检索日文文献提供了极大的方便．但如果这样做会产生什么问题呢？用什么去

替代那些日文中有而中文中无的字？对不符合中国人习惯而日文惯用的日文汉字词组怎么办，尽管这些词组中的每一个字都钶以认为是中文汉字．

我们以日本科学技术情报中心的内部汉字代码表为基础，即是以《速报》中可能使用到的全部汉字为基础，对日文科技文献，用数理统计的方法进行了大量的调查工作．研究表明，与中国汉字同字同意的日文汉字占其总数的绝大多数；字形有一些差别，但字意仍然相同的占少数；字形相同，字意不同或字形，字意完全不同的占绝少数．日文汉字的使用频率也呈相同变化趋势．统计见表 1．

表 1． 科技文献中日文汉字的特点

类号	字形	字意	比例	频度
1.	相同	相同	938‰	970‰
2.	不同	相同	58.7‰	20‰
3.	相同	不同	0.2‰	3‰
4.	不同	不同	3.1‰	7‰

据以上统计，第 1，2 两类占日文汉字的绝大多数，其使用频度也是最高．对这两类日文汉字采用以相应中国汉字直接替代的方式，字的笔划基本相同，语意密切接近，不但减少了烦锁的造字过程，变假名查询为拼音查询，同时也不会影响到检索结果的正确性．对第 3 类字形相同但字意不同的个别日文汉字，采取在查询过程中接口一个翻译对照表的方法加以解决．例如“信”，在对照表中的对应项是“手纸”．实际的检索用词即为“手纸”．最后，只对第 4 类日文汉字采取假名注音的方法加以解决．由于这部分日文汉字在整体上占很小很小的比例，其使用频度又很低，在查询过程中，查阅一下字典就成为一个微不中道的负担了．对词组的处理也是通过接口对照表的方法解决．例如中国人喜欢讲“控制”，而日本人则讲“制御”．但是输入人员如果直接输入符合日本民族习惯的词组的话，其查询时间将会大大减少．

通过以上处理，我们中国用户就能以我国文化为背景，按照我国的习惯避开语言障碍较为准确地检索日本科技文献．由于 MINISIS 的可开发性，使我们的人-机界面得到了很好的改善．

另外，为了在输入检索词时尽可能地简化操作，在 BATCHIN 前的转换程序中还进行了一些特殊处理．基本原则是：

1、对快速查询字段中能够以 ASCII 字符替代的 JICST 内部汉字代码一律替代，以便于键盘操作．

2、原带用双字符集制（EBCDIC HITACHI8000 系列和 JICST 内部汉字代码），经处理后，全部以 ASCII 代码和我所机内代码替代。并对组合字，黑体字，斜体字等标识符作相应处理，以适应 HP 输入输出系统。

由于程序采用了常驻内存，BIT 操作，递归调用和文件成块写入等措施，大大提高了数据的处理速度。转换程序的编制，还为后期处理提供了一些必要的词组和字段的分离与合并方面的能力。

以中文方式利用 MINISIS 检索日文文献，改善用户界面是本文的中心内容。在用 MINISIS 建立日本文献数据库的课题中，还得到过美国匹斯堡大学情报学和图书馆学系主任，安东尼・戴本斯博士的指导和帮助，在此表示感谢。

参考文献

1、MINISIS MANUALS （1）－（3）
2、HP3000 COMPUTER LANGUAGE & SUBSYSTEMS
3、HP CHINESE I／O SYSTEMS
4、INFORMATION RETRIEVAL SYSTEMS
CHARACTERISTICS TESTING AND EVALUATION
2ND EDITION F WILFRID LANGCASTER
5、FUNDAMENTALS OF CHINESE LANGUAGE COMPUTING
6、FIOS 活用の手引き （1）－（4）

Setting up Large Scale Document Searching Data Base with the MINISIS DBMS

LIU, YUN
The Scientech Information Centre
The Ministry of Machinery & Electronic Industries

MINISIS is a relational data base management system. It supplies many functions, such as data base definition, data base updating, inverted file updating, query, etc.. These functions are suitable for document searching. To set up data bases using data bases available on readable computer tapes, there's no need for you to extend MINISIS's functions. A few simple steps allows you to load data from these tapes into a MINISIS data base. First, compare the format of the data on the tapes with the format required by these MINISIS modules: BATCHIN, INDEX, INVERT. After you know the differences, modify the format of data on the source tapes until it matches the format required by the MINISIS software.

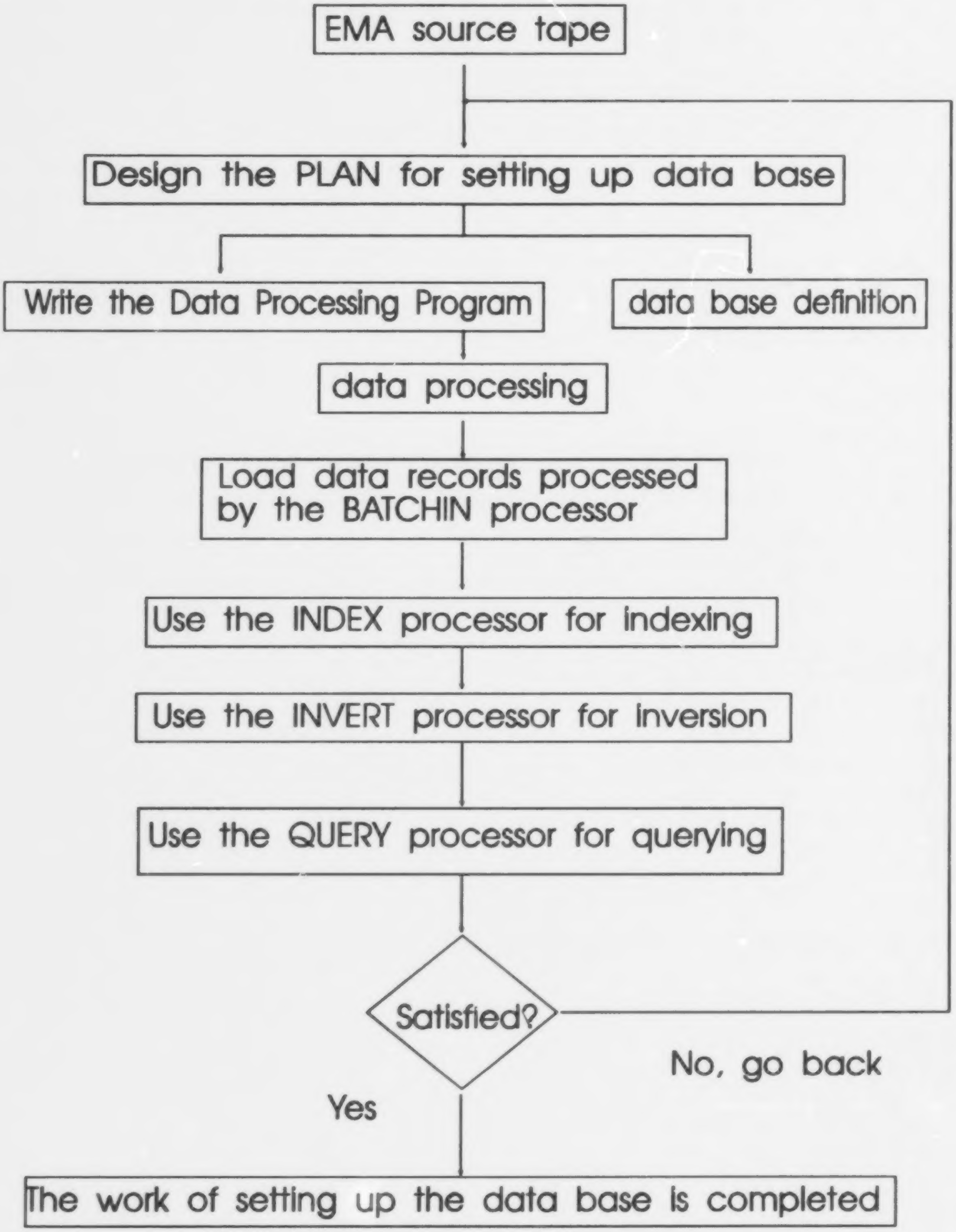
In this article, I'd like to introduce some of my work of setting up document searching data base by using the EMA (Engineered Material Abstracts) document tapes and the MINISIS software. I hope that it's something useful for you.

I. General ideals for setting up data base

EMA document tapes are published by the American Society for Metals (ASM). It's contents contain ceramics, composites, polymers and their usages. The data format of EMA data tapes is defined by the ASM, and is different from the format required by the MINISIS software. So before you can load data on the EMA tapes into the MINISIS data base, you must change the format of the EMA source tapes.

First we searched both the format of EMA data tapes and the format requested by the MINISIS carefully. We used PASCAL 3000 to write a Data Processing Program. We use this program to modify the data format of EMA. The data format changing is the most important part of the work. After that we can load the processed data records into the data base using the MINISIS BATCHIN processor. Tasks, such as indexing, inversion, query, etc. can all be done by the MINISIS's associated processors after the data is loaded.

The following is the flow chart in setting up the data base.


EMA source tape
Design the PLAN for setting up data base
Write the Data Processing Program
data base definition
data processing
Load data records processed by the BATCHIN processor
Use the INDEX processor for indexing
Use the INVERT processor for inversion
Use the QUERY processor for querying
Satisfied?
No, go back
Yes
The work of setting up the data base is completed

Here I'd like to point out two important aspects, 1) before I can do the work of setting up the data base I designed a plan for data processing. This plan plays the key role in all of this work. The Data Processing Program is written according to this plan, and the data base is also designed according to this plan. 2) The Data Processing Program is written using the structured design method. The main body of the program does basic operations, such as reading data blocks from the input file, character code translating from EBCDIC to ASCII, linking physical blocks to generate a complete logical record, writing processed records to the output file, etc. The other tasks, such as fields processing, is accomplished by the associated subroutines or procedures. The advantage of this program design method is that when you want to change the processing of any field, you only need to modify the associated sub-program, and the other part of the program can remain unchanged.

II. The processing of each field of EMA records

1. Document Number Field

This field consists of several contents, such as year, month, subject category, material designation, etc.. After processing, this field is separated into a group of subfields contained in one subfielded field, each content corresponds to one subfield. And this makes it possible to index from some of these subfields MINISIS's INDEX processor.

2. Title Field

This field contains the title of the bibliographic record. There are no changes in format for this field and in order to use the facility of adjacency searching supplied by the MINISIS's G.00 version, the ADJ type of B-tree inverted file is used for this field.

3. Abstract Field

This field contains the complete abstract text. There are no changes for this field and in order to save disk space, the field is not inverted.

4. Personal Authors and Corporate Authors Field

This field contains the names of the authors of an item. Personal authors are separated from corporate authors by "$". Within the personal authors and the corporate authors, each name is separated by a comma and a space. After processing, this field is separated into two fields – personal author field and corporate author field. Each of this two fields is repeatable and is inverted field.

5. Document Source Field

This field contains the bibliographic information pertaining to the item. It is composed of many contents separated from each other by a "$" and followed by a single letter denoting the next content. After processing, this field is separated into a group of fields and some of these fields are subfielded field.

6. Indexing Terms Field

This field contains controlled index headings assigned to the document from Engineered Materials Thesaurus. Each index entry consists of a primary term, which may be a single word or a multi-word phrase, and an optional subordinate term, which may also be a single word or a multi-word phrase. The subordinate term is separated from the primary term by a comma and a space; index entries are separated by a /'

After processing this field is changed to subfielded field. It contains two subfields – primary term subfield and subordinate term subfield. The subfielded field is repeatable, and each occurrence corresponds to an index entry. The advantages of such processing are that we can use the INDEX processor to index both the primary terms and subordinate terms and we can use the thesaurus facility of MINISIS for the primary term subfield.

7. Trade Names and Material Classification Field
This field contains specific manufacturer's designations or trade names of material(s) discussed in the source document, and a materials classification code identifying the family group to which the individual material(s) belongs. Trade names are separated from the related material classification code by a semi-colon, and within the trade names, each name is separated by a comma and a space. A record may contain several material classifications with associated trade names. Each trade name(s); material classification set is separated by a "$".

After processing, this field is changed to a subfielded field with two subfields – trade name subfield and material classification subfield. The subfielded field is repeatable, each occurrence corresponds to a trade name(s); material classification code set. So both of the trade names and material classification code can be indexed and inverted by the MINISIS's INDEX and INVERT processors.

III. Data Base Updating

The data base updating work involves the following: EMA source tape processing, batch loading for the processed data records, index, and inversion. The first part of the work is done by using the Data Processing Program written by the data base manager, others are done by MINISIS's related processors, such as BATCHIN, INDEX, INVERT, etc.. So as a data base manager, the most important part of this work is data processing. He must test his program carefully, look again and to find bugs which may exist in his program. He must be sure that his program is indeed correct. Then he can use his program to process the EMA source tapes, load processed records into data base, and do the other sub-works step by step, and the data base for document searching will be set up.

利用 MINISIS 数据库管理系统建立大型文献数据库

机械电子部机械科技情报所
机械中心图书馆软件部
刘云

1989 年 9 月 19 日

MINISIS 是一种关系型数据库管理系统,它具有方便灵活的数据定义、数据更新、倒排文件维护、查询等等功能,这些功能对于文献检索来讲基本上是完备的. 因此, 作为一个数据库经理, 当他需要利用 MINISIS 软件将某种现成的、已经存贮在磁带上的数据记录建成数据库的时候, 他主要应该做的并不是对 MINISIS 软件本身作许多扩展, 而应该是深入细致地研究这种记录在磁带上的数据格式和 MINISIS 软件所要求的数据格式, 找出两者之间的差异, 然后设法对磁带数据的格式进行修改,直到它和 MINISIS 软件所要求的数据格式完全一致为止. 格式经过了修改的数据记录被 MINISIS 软件以批式方式读人到数据库以后, 接下去的几乎全部的工作,诸如抽词、倒排、检索等等,都将由 MINISIS 软件来完成,而数据库经理所需要做的工作只是使用 MINISIS 的有关功能模块就行了.

本文想介绍一下笔者利用 MINISIS 软件将《工程材料文摘》(Engineered Material Abstracts, 简称 EMA)英文文献磁带建成大型联机数据库的工作情况, 并愿意将其中的一些经验提供给各位参考.

一. 数据库的总体构思

《工程材料文摘》(下简称 EMA)英文文献磁带是美国金属学会(American Society for Metals)出版发行的一种文献磁带,内容包括陶瓷、复合材料、高分子材料等非金属材料及其应用. 这种磁带的数据记录的格式是磁带出版发行机构自己定义的,它和 MINISIS 的批式读人模块 BATCHIN 及其它一些模块对数据格式的要求存在着很大的差别,这些差别归纳起来有以下四个方面: 1. 记录格式的差别——这种差别包括逻辑记录起止的标志、字段标识符的形式等,必须消除这种差别,磁带数据才能被 MINISIS 的 BATCHIN 模块所接收,也就是说,才能被装人到数据库中. 2. 字段安排上的差别——EMA 文献磁带只提供基本数据,为了著录的方便,形式上只设立了八个字段,但实际上往往是一个字段包含若干项内容,如果就这样装人数据库,许多字段将无法进行检索,而且也不便于用户阅读;为了使今后所建成的数据库有多个方便灵活的检索途径,以满足用户多方面的检索要求,以及使数据库有清晰明了的显示和打印格式,有必要把原磁带记录中的某些字段分拆成若干内容单一的字段,或者在原字段的基础

上派生出新的字段。3. 与抽词倒排要求的差别——原数据记录中的某些字段在形式上不符合 MINISIS 的 INDEX 模块对抽词的要求，这种差别将直接影响到以后的倒排的进行，因而必须加以消除。4. 逻辑记录与物理块对应关系的差别——在 EMA 磁带中，逻辑记录是按照每个物理块 1202 个字节记录在磁带上的，而 MINISIS 的 BATCHIN 模块要求批式读人数据的物理块长最多不得超过 256 个字节，因此，原磁带中逻辑记录和物理记录的对应关系也需要进行修改。综合以上所讲的四种差别，本人在实际开始建库工作之前，首先制定了一个详尽的“磁带数据处理方案”，该方案详细指明了应该如何对原磁带数据的格式进行修改；然后本人在 HP3000–68 计算机上用 Pascal 语言按照该方案编写出了磁带数据处理程序，把原磁带数据处理成完全符合 MINISIS 软件要求和文献检索要求的形式。接下来的一系列工作，诸如数据库模式的定义、批式读人数据、抽词、更新倒排文件、检索等等，都是利用 MINISIS 所提供的有关的功能模块来完成的。

下面是整个建库工作的流程图：

《工程材料文摘》磁带

制定“磁带数据处理方案”

编制数据处理程序　　　设计数据库模式

对磁带数据进行处理

批式装人数据库

抽词

倒排

检索

满足各项要求吗?　　　否，返回修改

是

正式装库

这里需要着重指出两点。第一点,“磁带数据处理方案”在整个建库工作中起着非常重要的作用,数据处理程序将严格按照这个方案来编写,数据库模式也将按照这个方案来设计;更重要的是,在建库工作尚未实际开始,即实际的数据库还根本不存在,甚至数据处理程序还没有编写,数据库模式也未设计的时候,从这个方案上就完全可以看到今后所建成数据库的全貌,大量的论证和修改工作都在这个方案上进行,等到这个方案已基本完善以后,才按照这个方案实际编写出数据处理程序、进行数据库模式的设计,依次完成各项后续工作。实践证明,这种先制定建库方案,然后才实际开始建库工作的工作方式,不仅大大提高了工作效率,缩短了数据库的研制时间,而且避免了因数据库建成后才发现有不合理的地方而不得不进行的大量反工,从而避免了大量人力、物力的浪费,在实际工作中收到了很好的效果。第二点,数据处理程序是用 Pascal 语言编写的,而且完全采用了结构化程序设计的思想,即主程序只完成某些基本的操作,例如从输入文件中读取数据、EBCDIC 码和 ASCⅡ码的转换、将各物理块链接成完整的逻辑记录、将处理好的逻辑记录写入输出文件等,这些操作基本上是固定的,因此主程序一经调试好以后就不会再有大的改动;而大量的修改、处理工作是由不同的子程序来完成的,子程序之间通过指针来发生联系,当某个字段的处理方案有改动的时候,只需要修改相应的子程序即可,对主程序和其它的子程序并不产生任何影响。实践证明,这样的程序设计方法不仅使整个数据处理程序条理清晰、层次分明、便于阅读和修改,而且给程序的调试带来了很大的方便,程序的正确性也容易得到保证。

二、EMA 文献磁带中各字段的处理情况

1. 文献号字段(Document Nunber)

该字段包括磁带出版的年份和月份、主题索引、材料粗分类码、流水号等五项内容。在处理中我们将该字段按照内容分拆成了一个母字段下的五个子字段,其中年份、主题索引、材料粗分类码三个子字段作为倒排字段,可进行快速查找。

2. 标题字段(Title)

该字段是倒排字段．为了便于对自然语言进行文中查找，倒排时利用了 MINISIS 的 G 版本所提供的 ADJ 型倒排文件，可进行相邻检索．

3. 文摘字段(Abstract)

该字段保留了原来的形式，末作修改．而且由于文摘字段一般都较长，为了节省磁盘空间，该字段没有倒排，可进行顺序查找．

4. 个人作者和团体作者字段 (Personal Authors and Corporate Authors)

在 EMA 磁带中，个人作者和团体作者是放在同一个字段中，中间用"$"符号分开，这样不便于进行倒排和检索，所以处理后将个人作者和团体作者分开作为两个字段，这两个字段都是重复字段，而且都进行倒排．

5. 文献来源字段(Document Source)

在 EMA 磁带中，该字段的结构最为复杂，共包括 17 项内容，即会议名称、会议地点、会议日期；出版者姓名、出版地点、出版日期；刊名、语种、卷期号、页码等等．这么多项内容放在一个字段里，既不便于检索也不便于阅读，所以处理后将该字段分拆成了若干个字段，有的字段还包含若干个子字段，以便把关系较为密切的项在逻辑上放在一起．

6.索引词字段(Indexing Terms)

在 EMA 磁带中，是把全部索引词放在一个字段里，词与词之间用"$"符号分开；而同一个索引词又可能是由主标题词和副标题词两部分组成，主、副标题词之间用"，"分开．这样的字段安排不便于 MINISIS 软件抽词、倒排的进行，也不便于使用 MINISIS 软件所提供的词表功能，因此处理以后将该字段变成了带子字段的重复字段，母字段的一次出现对应于一个索引词，而主标题词和副标题词分别放在同一个母字段的两个子字段中．这样的处理方式不仅在逻辑上仍然保持了原磁带中主、付标题词的配对关系，而且使得 MINISIS 软件所提供的抽词、倒排和词表功能能够被顺利使用．

7.商品名称和材料分类码字段(Trade Names and Material Classification Code)

在 EMA 磁带中，该字段由若干个组构成，组与组之间用"$"符号分开；在同一个组里，包含一个或多个商品名称以及这些商品名称所对应的材料分类码．为了抽词倒排和检索的需要，我们将该字段处理成带子字段的重复字段，每一个组作为重复字段的一次出现，在每一次出现中又有两个子字段，即商品名称字段和材料分类码字段，这两个子字段都有检索意义，因而都进行了倒排．

8.相关参照及保密码字段(Cross Reference and Security Code)

在原磁带中，该字段包括相关参照码和保密代码两项内容，中间用"$"符号分开．虽然这两项内容并没有检索意义，不需要进行抽词和倒排，但为了便于今后有其它用途，因而处理后仍然将这两项内容分折成两个子字段；但在显示和打印输出时，又将这两个子字段合在了一起，因而这种字段的分拆对于终端用户界面来说是看不到的．

三.数据库的维护和更新

“工程材料文摘”英文文献数据库是利用 EMA 文献磁带建成的,随着磁带出版发行机构不断地提供新的磁带,需要对数据库不断地进行维护和更新．整个数据库更新工作由以下四部分组成,即原磁带数据的前期处理、数据批式装库、抽词、倒排．其中第一部分工作是用数据库经理自己编写的程序完成的．而后面的三部分工作都是由 MINISIS 软件所提供的有关的功能模块来完成的．因此数据库经理需要把大部分注意力放在第一部分工作上．因为在数据处理程序投入运行的初期．常常会遇到数据中出现一些意想不到的情况．使得数据处理程序的运行中途停止．这时的数据处理程序是处在一个不稳定期．这就需要在数据处理程序中预先设置一些“断点”．以指示出错的部位．并指示出是哪一条数据记录使得处理程序出错．以便对处理程序进行相应的修改．等到数据处理程序的运行基本上稳定以后．整个数据库的维护和更新工作就可以交给专门的操作人员来完成,这时对于数据库经理来讲,整个建库工作基本上就结束了．

四.结束语

本人在研究和建立“工程材料文摘”英文文献数据库的过程中,得到了机械中心图书馆各位馆领导的热情关怀和大力支持,本数据库的建成,是与各位领导所创造的各方面的有利条件分不开的,在此表示衷心的感谢!本所高级工程师张晰老师、吴亢老师、杨则正老师和软件部的李林甲老师为数据库的建立给予了许多宝贵的指导．在此一并表示感谢．

Go Further into the Retrieval Capability of Specialized Data Base

Wu Bing
The Scientech Information Centre
The Ministry of Machinery & Electronics Industries

People who retrieve information through computers are interested in how to find the required information in a simply and accurately. The data base management system can meet the retrieval requirements if the data are not very complicated. It is not easy to fulfil strict retrieval requirements when the data are very complicated.

METADEX is an abstract journal published by the American Metallurgical Society and the British Metallurgical Society. We used machine readable METADEX to create our METADEX data base. The METADEX data base in the Scien-Tech Information Centre was created in 1986 using the MINISIS data base management system.

This is an example of a record in the METADEX data base:

```
ISN 320          PY 1985
AN  850113-0041
TI  Aspects of the Microsegregation of Impurity and
    Alloying Elements in Alpha-Titanium Alloys
AU  S.S. Ushkov et al.
JN  Fiz. Met. Metalloved.   Jan. 1984, 57, (1),
PG  194-197
IN  ISSN 0015-3230
LA  Russian
AB  Cf. Kryukov et al., ibid., 1981, 52, 880.
    Grain-boundary segregation has been investigated by
    X-ray spectral microanalysis in: (1) iodide Ti, (2)
    commercial Ti, (3) Ti with 2.0% Al and 2.5% Zr and
    (4) Ti with 5.5% Al, 6% Zr and 1.5% V. The results
    show that in alloy (4) there is a sharp fall in the
    segregation of Fe and Ni at the grain boundaries.
    This is due to the fact that in the presence of Al
    the activation energy considerably increases, while
    the diffusion coeff. of Fe and similar elements in
    alpha -Ti decreases. This explains the smaller size
    and number of particles precipitated at the grain
    boundaries as a result of prolonged ageing in the
    multicomponent alloy (4) as compared with Ti and
    alloy (3). Thus one of the aspects of multicomponent
    alloying is that diffusion movement of addition or
    impurity atoms is retarded and damage at the grain
    boundaries is reduced.--N.V
DE  Titanium base alloys, Crystal lattices/ Aluminum,
    Alloying elements/ Transition metals, Alloying
    elements/ Segregations/ Diffusion, Alloying effects/
    Aging (artificial)/
ST  TI;VTi-0/ TI,Ti-13.49Zr-16.45V-6.35Fe-0.88Ni/
    TI;Ti-39.51Fe-16.68Ni-1.89Zr/
CC  13
```

The alloy name field needs high quality retrieval access. The ST field is the alloy index field. The alloy index is the main field for retrieval from the METADEX data base. The underlined words are the alloy name. The alloy name is abbreviated using letters adopted from the English alphabet. The numerals show the element's proportion. The hyphen divides the elements. If a document mentions more than one alloy name, each of the alloy names in the field are arranged sequentially. It is not difficult to retrieve an alloy name, if one knows the alloy name exactly as it is indexed. From the view of the METADEX data base users, the following problems should not be overlooked:

1) Most of the users hope to find some documents for reference.
2) A user should not have to guess at the sequence of the elements.
3) Users should not have to know all the non-alphanumeric characters in the alloy names, such as hyphen "–."

They may not know the exact indexing alloy name before searching. For example, a user expecting a reference document on an alloy including titanium, iron, nickel, and zirconium would have to know the proportion of the elements to retrieve a document indexed with an alloy name Ti-39.51Fe-16.68Ni-1.89Zr. Perhaps a document about the same alloy in which the proportion of Iron is 39.41 is useful too!

From these problems we can see that although there is an inverted file for alloy names, it is almost impossible to retrieve by alloy name unless one knows exactly how the alloy was indexed. The alloy name inverted file is obviously created for show. The employment of this access path is very low.

In the United States there is a METADEX data base in the authoritative DIALOG Information Retrieval System. How did DIALOG System solve the alloy indexing problem? In the DIALOG system, the alloy name filed of METADEX data base can be retrieved by its exact indexing alloy name and it can be also retrieved through the element symbols of the alloy. This could create some errors when there is more than one alloy indexed for one document.

For example, if a document is indexed by two alloy names: Al–41.2Ni and Al-1.23Zr and a user searches for an alloy which contains nickel and zirconium using Ni and Zr. The document indexed by Al-41.2Ni and Al-1.23Zr is retrieved. This hit is a mistake because although the elements Ni and Zr occur in the same document, they do not occur in the same alloy name.

No functions provided by MINISIS can fulfil these retrieval requirements directly. Further developments enhanced the retrieval capability of the alloy name field, and data processing improved. MINISIS functions were used nimbly. After successful development, the retrieval requirements of alloy name field can be fulfilled as expected by the users. In the new development, the repeatable alloy name field was flattened and loaded into the data base. The following retrieval methods can be used after the improvement.

For example, the alloy names for document A are:
Ti-13.49Zr-16.45V-6.35Fe-0.88Hi and Ti-39.51Fe-16.68Hi-1.89Zr
Document A can be retrieved using one of these methods:
Retrieval method 1) Retrieve by element symbols within the alloy names, such as:
Q ST Ti and Fe and Hi and Zr.
Retrieval method 2) Retrieve by element symbols within the alloy names regardless of the sequence of occurrence for each element. such as:
Q STV and Hi and Fe and Zr and Ti.
Retrieval method 3) Retrieve by element symbols within the alloy names. But some of the elements should not be entered, such as:
Q ST Ti and Fe.
Retrieval method 4) Retrieve by using the exact indexing alloy name, such as:
Q ST Ti-39.51Fe-16.68Hi-1.89Zr.

From the retrieval methods described above, we can see that when retrieving the alloy name field, a data base user doesn't have to pay attention to the special characters in the alloy name, the element's proportion and the sequence of occurrence for each element. This new function of METADEX data base in Scien-Tech Information Centre of the Ministry of Machinery and Electronic Industry of China is much better than that of the METADEX data base in the DIALOG system of the United States. The new development provides a powerful retrieval method for the metallurgists.

Although data and data base management systems are similar, data base design and data processing affects results. There is still a lot to do on designing powerful data bases under the existing data base management systems.

References:

1. MINISIS Manuals

2. DIALOG Information Retrieval Manuals

对提高专业数据库检索功能的深入探索

吴 冰

使用计算机检索资料的人最关心如何用简便的方法准确地查到所需资料．数据内容比较简单的数据库能够直接利用数据库管理系统的功能去满足检索的要求．但是对于数据内容比较复杂、而且对检索要求很高的数据库来说，达到满意的检索要求不是一件很容易的事．

《金属文摘》是美国金属学会和英国金属学会出版的检索刊物．我们建立《金属文摘》数据库使用的是机器可读版本．《金属文摘》数据库是于1986年用MINISIS数据库管理系统建立的．数据库中合金牌号字段是检索要求比较高的数据项．下面列出《金属文摘》的一篇记录．

```
ISN 320         PY 1985
AN  850113-0041
TI  Aspects of the Microsegregation of Impurity and
    Alloying Elements in Alpha-Titanium Alloys
AU  S.S. Ushkov et al.
JN  Fiz. Met. Metalloved.   Jan. 1984,  57, (1),
PG  194-197
IN  ISSN 0015-3230
LA  Russian
AB  Cf. Kryukov et al., ibid. , 1981, 52, 880.
    Grain-boundary segregation has been investigated by
    X-ray spectral microanalysis in: (1) iodide Ti, (2)
    commercial Ti, (3) Ti with 2.0% Al and 2.5% Zr and
    (4) Ti with 5.5% Al, 6% Zr and 1.5% V. The results
    show that in alloy (4) there is a sharp fall in the
    segregation of Fe and Ni at the grain boundaries.
    This is due to the fact that in the presence of Al
    the activation energy considerably increases, while
    the diffusion coeff. of Fe and similar elements in
    alpha -Ti decreases. This explains the smaller size
    and number of particles precipitated at the grain
    boundaries as a result of prolonged ageing in the
    multicomponent alloy (4) as compared with Ti and
    alloy (3). Thus one of the aspects of multicomponent
    alloying is that diffusion movement of addition or
    impurity atoms is retarded and damage at the grain
    boundaries is reduced.--N.V
DE  Titanium base alloys, Crystal lattices/ Aluminum,
    Alloying elements/ Transition metals, Alloying
    elements/ Segregations/ Diffusion, Alloying effects/
    Aging (artificial)/
ST  TI;VT1-0/ TI;Ti-13.49Zr-16.45V-6.35Fe-0.88Ni/
    TI;Ti-39.51Fe-16.68Ni-1.89Zr/
CC  13
```

ST 字段是合金索引字段．合金索引是《金属文摘》中一项重要的检索途径．用下划线标出部分是合金牌号．合金牌号中的字母是合金元素的缩写名，数字部分代表合金元素所占的比例，短横“–”将每个元素分开．合金索引字段中所标出的合金牌号都是该篇文献中所涉及到的合金．如果一篇文献中涉及的合金牌号不止一种，则合金索引中的合金牌号标引也重复多次．检索时，如果完全按照文摘中所标引的合金牌号去检索是很容易办到的．但是设身处地从数据库用户角度来考虑一下检索合金牌号的问题：

1）大多数用户是查找有参考价值的文献．他不大可能事先知道完全准确的合金标引方式．比如 Ti–39.51Fe–16.68Ni–1.89Zr 是某篇文献的合金牌号标引，用户也要查含有钛、铁、镍、锆这四种元素的合金．但是用户不可能事先知道铁元素的比例是 39.51．也许含铁量比例为 39.41 的这种合金也很有参考价值呢．

2）合金牌号中，每种合金元素出现的次序也很难由用户来推测．

3）合金牌号中会出现一些既不是字母、也不是数字的特殊字符，如：“–”也很难事先想象出来．

从上面问题可知：既使为合金牌号建立倒排文件，除非事先知道准确的合金牌号的标引方式，否则检索合金牌号几乎是不可能的．这样，倒排的合金牌号文件只是摆摆样子而子，利用率很低．

很有权威的美国 Dialog 情报检索系统也有《金属文摘》数据库．Dialog 是怎样解决这个问题的呢？Dialog 系统除了提供完全按合金牌号检索的功能外，还提供按合金所含元素的缩写名检索的功能．但是当一篇文献标引多个合金牌号时，用合金中元素的缩写名检索出现误检的可能性很大．例如某篇文献标引了两个合金牌号：Al–41.2Ni 和 Al–1.23Zr．而当用户用 Ni 和 Zr 去检索某种合金时，这篇文摘也能命中．但是命中文献不是用户所需要的．这种命中是错误的．因为 Ni 和 Zr 元素虽然出现在同一篇文摘中，但是不同时出现同一合金牌号中．

在 MINISIS 提供的功能中，找不到一种功能够直接满足上述检索要求．但是为了改善《金属文摘》数据库检索功能，灵巧地运用了 MINISIS 现有功能，并对数据进行了进一步处理．使得合金牌号字段的检索功能完全达到了预想的效果．我们采取了将可重复的合金牌号字段拉平后再装入数据库的方法．

经过处理的合金牌号字段可以用以下方式进行检索。

如，一篇文献中的合金名为：

Ti–13.49Zr–16.45V–6.35Fe–0.8Ni　和

Ti–39.51Fe–16.68Ni–1.89Zr

检索方法 1、用元素缩写名检索，可命中。如：

ST　Ti and Fe and Ni and Zr

其中 ST 是合金牌号字段的记忆名。

检索方法 2、用元素缩写名检索，不考虑元素在合金牌号出现的顺序，可命中。如：

ST　V and Ni and Fe and Zr and Ti

检索方法 3、用元素缩写名检索，只输入合金牌号中的部分元素缩写名，可命中。如：

ST　Ti and Fe

检索方法 4、完全按照合金牌号检索，可命中。如：

ST　Ti–39.51 Fe–16.68Ni–1.89Zr

从上述检索途径可知，用户检索合金牌号时，完全不用考虑合金牌号中出现的特殊字符，不用考虑合金元素所占的比例系数，不用考虑合金出现的顺序。机电部机械科技情报所建立的《金属文摘》数据库的这项检索功能完全超过了美国 Dialog 系统中《金属文摘》数据库合金牌号的检索功能。这项检索功能的开发为从事合金研究的工作者提供了有力的检索手段。

实践证明，同样的数据，同一个 MINISIS 数据管理系统，使用不同的数据库设计方法和数据处理方法可以得到相差很大的结果。由此可见，如何利用现成的数据库管理软件设计出一个好的、功能强的数据库是大有文章可作的。

参 考 文 献

1. MINISIS 使用手册
2. METADEX 磁带使用说明
3. Dialog 联机情报检索手册
4. 《金属文摘》主题词表

THE APPLICATION OF MINISIS IN UNION CATALOG OF PERIODICALS

Wang Yan-mei
Library of the Chinese Academy of Medical Sciences (CAMS) and Peking Union Medical College (PUMC)
Translated into English by Chen Cheng-pei

[Abstract] This article introduces the general state of two data bases of the "Union catalog of Periodicals". The audience served, objectives of the service, special features, and the application of the catalog and the manner used to process various data bases are discussed. The important role of processors such as PRINT, BACHIN, INDEX and QUERY are explained whereas the data transferring process between the MINISIS data base and the dBASE III base is identified.

I. GENERAL CONDITION

The Library of the CAMS and PUMC recently introduced the automated management of periodicals. We organized the editing of two union catalogs of biomedical periodicals, and then used these for the management of periodical work. Now we have two data bases of union catalog of periodicals. They are different from each other in the audience served, special features of service, their data base definitions, and processing methods.

II. Data base of the "Union Catalogue of Periodicals in Western Languages Subscribed by Libraries of Medical and Medicinal Universities and Colleges in Whole China". (UCPS)

1. The audience served and special features of service of the UCPS

The audience served by the UCPS comprises most member libraries of the "Association of Libraries of Chinese Medical and Medicinal Universities and Colleges", numbering 111. The annual catalog serves the working staff of the libraries. It provides information periodical subscriptions, and resource sharing and coordinating.

2. Characteristics of the Data base and its Application

1) Define features of the data base

a. Basic condition of data base definition:

TYPE	MAX SIZE OF MASTER FILE	AUTO-NUMBERING	MAX NUMBER OF ISN	NUMBER OF FIELDS	NUMBER OF INVERT FILES
RD	5000	N	5000	4	3

The data base takes the title of every single periodical as the mark of one item of record. Now there are more than 2,500 items of record in the data base.

b. Basic condition of field definition is as follows:

FIELD NAME	LENGTH	REPEATABLE	NUMBER DATA TYPE	MANDATORY FIELD	INVERT
TITLE	250	N	N	Y	Y(ONL)
CHINESE SERIAL NUMBER	12	N	N	N	Y(ONL)
LIBRARY CODE	4	Y	N	Y	Y(ONL)
NUMERAL	2	N	Y	Y	N

2) The application of data base
The data base was established in Autumn 1987. It was used in the UCPS catalog plate making process and the retrieval of data. In the years 1988 and 1989 services to users were provided. In the beginning of every year the catalog is published in book form. Beginning from this year (1989), diskette data of the catalog will also be available.

The data base provides three retrieval entry points, "TITLE", "CHINESE SERIAL NUMBER" and "LIBRARY CODE". The "NUMERAL" field is used for statistical calculations. It is a mandatory field accepting only the numeral "1" as input. If we want to estimate the number of kinds of periodicals a certain unit subscribed, we must conduct conditional retrieval in COMPUTE processor and accumulate the numbers of the record with conditions qualified at the same time, i.e. accumulate the "NUMERAL" field after performing a relational operation.

By operating GARBAGE, we have got some valuable data, including the longest record length, average record length, records that contain most field numbers, average numbers, blank space ISN in data base etc. These data have facilitated the later-on data output, plate-making, the estimation of how many pages the catalog needed, record appending or deleting and so on.

At the beginning of this year, after we made an approach to the data transfer between MINISIS and dBASE III we have more succesfully put the data of two years, 1988 and 1989 onto diskette. We supply the catalog on diskette to each of the Chinese libraries with the appropriate level of automation. This is also an important aspect showing the result of the application of MINISIS data base. This task is done using the PRINT processor. We output data to text files formatted for input to dBASE III. Quite a lot of work has been done in the "Literal" section of the PRINT field attribute definition. To answer the following 4 point correctly:

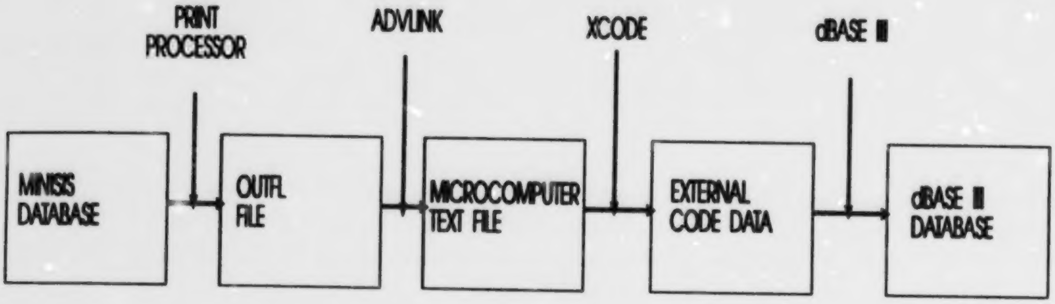
LITERAL CONDITIONAL – –

LITERAL POST – –

LITERAL REPEATABLE – –

LITERAL SUPPRESSION – –

The aim is to add quotation marks before and after each field, to add comma between the field and the answer to CONDITIONAL DISPLAY is NO. The data processing process from MINISIS to dBASE III can be illustrated as follows:



3) The maintenance of the Data base
The publication cycle of the UCPS catalog and the service effectiveness usually one year demand strict timeliness. The collecting and modifying of data and the accuracy of the UCPS requires hard work. Each year we send the member libraries a "notice to collect". Every two years we send the list of their periodicals catalog asking that they check and correct it. The INDEX and the PRINT processor of MINISIS play an important role in this work. Proper pre-processing work is to be performed upon the collected data. Using the MPE VPLUS subsystem we come online to modify the original data base record. After this a new UCPS catalog can be output. Meanwhile statistic data are obtained, which could hardly be accomplished by manual labour. For instance, compute the numbers and kinds of periodicals the 111 libraries subscribe to or the most and the least kinds of periodicals each library subscribed to.

The need for maintenance also applies to BTREE files. To modify the data directly affects the dynamic BTREE space. When necessary, it is imperative to utilize the TREEMANT function in INVERT processor to expand BTREE, ensuring accuracy, reliability and availability to the data in data base.

This data base is characterized by its being concise and exquisite as well as its data continuity. Under the operation of computer processing, its timeliness and accuracy serves the vast periodical management personnel and the readers.

III. The Data base of the "Union Catalog of Biomedical Periodicals in Foreign Languages in Whole China". (UCBP)

1. The audience served and special features of service of the UCBP

There is a little difference between the audiences of the UCBP and the UCPS. The former has been extended to include libraries of scientific research units or institutes. The collecting and editing of the UCBP Catalog is done once every 5 years. Its main objective is to serve the reading public. Its provides convenience of resources sharing and necessary data for the accumulation of resources materials.

2. Characterizations of the data base and the processing process of data

1) Definition characteristics of data base

a. Basic condition of data base definition:

TYPE	MAX SIXE OF MASTER FILE	AUTO-NUMBERING	MAX SIZE OF ISN	NUMBER OF FIELDS	NUMBER OF INVERTED FILES
RD	5000	N	8000	21	11

In the past the editing of the UCBP was accomplished entirely through manual labour, and was treated as a one-time task without continuity of data. Automation not only saves man power and time, but also improves the compilation of the UCBP catalog. Some of the main features have been reflected as follows:

Output Format of UCBP

```
ACM Transactions on Programming Languages
  and Systems . --v.1(1979). --New York,
  USA : Association for Computing
  Machinery, 1979
  eng. --Quarterly
  ISSN 0164-0925                   7388033

A.I.Ch.E. Journal : Chemical Enginneering
  Research and Development. --v.1(1955).
  --New York, USA : American Institute of
  Chemical Engineering, 1955
  eng. --Bimonthly
  ISSN 0001-1541                   8108001
```

```
AIDS Research : Acquires Immunodeficiency
  Syndrome. --v.1(1983). --New York, USA
  : Mary Ann Liebert,Inc., 1983
  eng. --Bimonthly
  ISSN 0737-6006                   6318029

AJNR : American Journal of Neuroradiology
  /American Society of Neuroradiology.
  --v.5(1984). --Baltimore, USA :
  American Roentgen Ray Society, 1984
  eng. --Monthly. --Continues: American
  Journal of Neuroradiology
  ISSN 0195-6108                   6358008
```

FIELD NAME	LENGTH	REPEATABLE	NUMBER DATA TYPE	MANDATORY FIELD	INVERT
TITLE	250	N	N	Y	Y(ONL)
SUBTITLE	250	N	N	N	Y
PARALLEL TITLE	250	Y	N	N	Y
STATEMENTS OF RESPONSIBILITY	200	N	N	N	N
YEAR ISSUE	30	N	N	N	N
PLACE OF PUBLICATION	50	N	N	N	N
PUBLISHER	150	N	N	N	Y
COUNTRY	4	N	N	N	Y
YEAR OF PUBLICATION	20	N	N	N	N
LANGUAGE	4	N	N	N	Y
FREQUENCY	50	N	N	N	N
NOTE	250	N	N	N	N
ISSN	12	N	N	N	Y
CHINESE SERIAL NUMBER	16	N	N	N	Y
CLASSIFICATION	10	Y	N	N	Y(ONL)
SUBJECT HEADING	80	Y	N	N	Y(ONL)
LIBRARY COLLECTION	60	Y (SUBFIELD FIELD INDICATOR)			
LIBRARY CODE	6	N	N	N	Y
VOLUME ISSUE	20	N	N	N	N
YEAR	30	N	N	N	N
NUMERAL	2	N	Y	Y	N

Note: The "Title" field is the default QUERY field.

– Subject heading and classification indexes: The addition of these two indexes has been long desired in the UCBP catalog but we were unable to do this before automation.

– Improved output using the PRINT processor: The manner of output keeps up with internationally standarized formats. Marks such as :,.,/,.–,(,), = ,etc. are automatically inserted by the PRINT processor.

– Using the INDEX processor to sort the data: The problem of misplacement is solved and results in an improved utilization rate of the UCBP and member libraries.

– Improved production value: Output using laser jet printers has improved the quality of plates.

2) To set up the data base and data processing

i) Pre-processing work – How this work is performed directly influences the quality of the data base. If the work is done carefully, regularly, the input and maintenance of data will be quick and smooth and the accuracy of the data can be guaranteed. Otherwise, it will bring much trouble to the proof-reader and maintenance personnel and confusion of data will emerge. The pre-processing work of this Catalog is done by librarians. To record subject heading and classification number is technical so it is done by associate research librarians.

ii) To set up the data base – The establishment of the UCBP data base is quite different from that of other data bases. The chief difference lies in the process of data transfer from the microcomputer to the minicomputer. The software manager sets up a temporary data base using dBASE III on the microcomputer for the input of data. MINISIS is used for the final data base.

This way the state of the host computer (whether the HP 3000 is operating or not, or the number of terminals available, or the availability of the computer management personnel) does not affect the work of the data entry personnel. This way, also, it is easy to achieve screen format input on the microcomputer. Version G was not available when this plan was developed. In comparing with non-Chinese terminals the superiority of Chinese prompting is more manifest. Finally, to avoid repetition and mistakes in input, parts of the fields can be replaced by code numbers. These are then replaced when the input goes through the intermediate routine (BASIC program) to produce BATCHIN format file.

iii) The process of input, transfer and further modification

The input work was completed by the special UCBP group of this library. Then the routine compilation was done by the software manager. After further processing, the data was output in the file format used by BATCHIN. All the data were put into the MINISIS data base. Nearly 3,100 items of record took about three hours. Further modification of the data involved the systematic analysis and check of the distribution of the contents of the 2 new fields: subjects and classification. Using the INDEX processor, they were separately designated as the primary keys, sorted and flattened, then titles were taken as secondary keys to thoroughly investigate the accuracy and appropriateness of the given subject headings and classification number. This point has shown the role played by MINISIS.

The data base being prepared for the UCBP catalog of 1986-1990 will be published in 1991. Except for the library collection field, all the fields have been input, verified and modified. After the collecting, arranging and inputting of library collection data in 1991, a new edition of UCBP will come out.

The existing data of the UCBP data base have laid a solid foundation for the forthcoming automation management of periodicals.

IV. A Brief Introduction to Data base of Resources Distribution Investigation of Biomedical, Periodicals in Whole China

This data base is very much the same as the above-mentioned two. The audience served is subscribers of biomedical periodicals throughout the country. It is not limited to libraries and institutions. The data is provided by CNPIEC (China National Publication Import and Export Company). This is once only task of resources distribution. Annual investigative statistics are made for them. Its processing manner is similar to that of the UCBP. The data base is set up with 14 fields and nearly 4,000 items of records. Now the data have been all input the microcomputer, are undergoing the transfer stage and will be input into the MINISIS data base very soon.

Concluding Remarks

Library automation is to a great extent simplified by the application of computers. At present, this task has been actively developing by our library. The editing of UCBP and UCPS will lay the foundation for the automation management system in this library. The acquisition and cataloguing system of books has come to an epilogue and the setting-up of Chinese data base of medical literature is under way. After the equipment has been augmented, an integrated management system in this library is going to be formulated and steadily improved by and by.

MINISIS 在外文期刊联合目录中的应用

中国医学科学院、中国协和医科大学图书馆
王燕美

【摘要】 本文主要介绍了两个期刊联合目录数据库的情况，从目录的收编范围、服务对象、数据库特征及其应用，以及各数据库的机处理方式等方面进行了论述。说明 PRINT、BATCHIN、INDEX、QUERY 等模块的重要作用。同时体现了 MINISIS 与 dBASE III 之间的数据传送过程。

[Abstract] This article mainly introduces the general state of two databases of the "Union Catalog of Periodicals". The covered scope, object of service, special features of the databases and their application of the catalog as well as the manner of computer processing of various databases are discussed. The important role of processors such as PRINT, BATCHIN, INDEX and QUERY are explained whereas the data transferring process between the MINISIS database and the dBASE III database reflected.

一. 概述

中国医学科学院图书馆的期刊自动化管理工作刚刚开始。我们首先着手的是组织编辑全国外文生物医学期刊联合目录，进而转为本馆的期刊管理工作。目前我们已建立了两个期刊联合目录数据库。它们均由西、日、俄等多文种组成，目前西文数据由计算机处理。两个目录在收编范围、服务特点，数据库在定义特征、机处理方式等方面各有所不同，以下进行分述。

二. 全国高等医药院校图书馆预订西文期刊联合目录数据库 （预订目录）

1.《预订目录》的收编范围及服务特点:

《预订目录》的收编范围是″中国高等医药院校图书馆协会″成员馆。现参加单位 111个。目录每年出版一次，以对图书馆工作人员服务为主，为每年的期刊预订工作及资源协调提供参考。

2.数据库特征及其应用

1)数据库定义特征:

a。数据库定义的基本情况:

TYPE	MAX.SIZE OF MASTER FILE	AUTO-NUMBERING	MAX.NUMBER OF SIN	NUMBER OF FIELD	NUMBER OF INVERT FILE
RD	5000	N	5000	4	3

数据库是以每一种期刊名称作为一条记录的标志。目前数据库中共有2500余条记录。

b. 字段定义的基本情况:

FIELD NAME	LENGTH	REPEATABLE	NUMBER DATA TYPE	MANDATORY FIELD	INVERT
TITLE	250	N	N	Y	Y (ONL)
CHINESE SERIAL NUMBER	12	N	N	N	Y (ONL)
LIBRARY CODE	4	Y	N	Y	Y (ONL)
NUMERAL	2	N	Y	Y	N

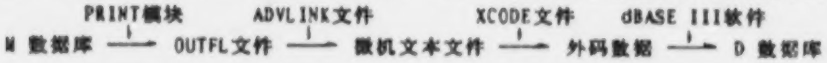
2)数据库的应用

数据库1987年建立，主要应用在为联合目录制版和数据的检索方面。88、89两年已向用户提供服务。每年初出版一期书本式目录，从今年起同时提供目录的微机软盘数据。数据库为用户提供了三个检索点，可以从刊名、刊号、馆代号三个字段分别检索，"数字"字段是用来作统计计算用的，是强制字段，且只能输入数字"1"。例如计算某单位订刊种数和占期刊总数的百分比，则需要在 COMPUTER 模块中，在进行条件检索的同时，逐个累加满足条件的记录个数，即累加"数字"字段，进行某些算术、关系运算之后，得到相应数据。通过对数据库做 GARBAGE，得到了一些很有价值的数据，主要有记录最大长度、平均长度，记录含最多字段个数、平均个数，数据库空白 ISN 号等。这些数据对后来的数据输出、排版、估算目录页数，对记录进行增删等，都提供了方便。

今年初，我们在进行了 MINISIS与 dBASE III之间数据传送工作的探讨之后，比较成功地将 88、89 两年的数据套录到微机软盘上。根据目前国内医药院校图书馆自动化水平，我们可以向一些单位提供软盘式目录。这也是MINISIS数据库应用的一个重要方面。这项工作的实施，主要是通过 PRINT模块，输出适用于 dBASE III 的数据文本文件。具体是在 PRINT 模块中定义字段属性时，在"文字"(LITERAL)说明部分下一点功夫。即对以下四点做出正确回答:

LITERAL CONDITIONAL ———，　LITERAL POST ———，
LITERAL REPEATABLE ———，　LITERAL SUPPRESSION ———，

目的是为每一字段前后加上引号，字段间加上逗号，对 CONDITIONAL DISPLAY 的回答是 NO。从 MINISIS 到 dBASE III 的数据处理过程可以表示为:



3)数据库的维护

由于《预订目录》的出版周期和服务效应都有着很强的时间性，因此在数据的征集、修改和保证其准确性方面给我们提出了更高的要求。每年向参加单位发一次征集通知，每两年发放一次各单位的上一期入编期刊名录，以供各单位进行核对和增删。MINISIS 的 INDEX、PRINT模块在此发挥了重要作用。在此基础上，将回收的数据做适当的前处理工作，借助 VPLUS子系统，对原有数据库记录进行全屏联机修改。校对之后，便可以输出新的一年的联合目录。同时得到手工难以完成的统计数据。例如：分别计算、输出 111个单位期刊的参编种数，订购最多、最少的期刊种类等。

数据库的维护也包括对 BTREE FILE 的维护。对数据的修改，直接影响到动太的 BTREE 空间。必要时， 要使用 INVERT 模块中的 TREEMANT 功能扩充 BTREE，确保数据库数据的准确、可靠和应用。

本数据库具有简而精的特点，也具有它的数据延续性。在运用计算机管理的前提下，以它的及时性和准确性服务于广大期刊管理人员和读者。

三．全国外文生物医学期刊联合目录数据库（联合目录）

1．《联合目录》的收编范围及服务特点

《联合目录》与《预订目录》的收编范围稍有不同，范围扩大到中国高等医药院校及科研单位图书馆、所。《联合目录》每五年收编一次，服务对象主要是读者。为资源共享提供方便，为积累资源资料提供数据。

2．数据库特征及数据的处理过程

1）数据库定义特征

a．数据库定义的基本情况（其中有 3 个联机倒排字段）

TYPE	MAX.SIZE OF MASTER FILE	AUTO-NUMBERING	MAX.NUMBER OF SIN	NUMBER OF FIELD	NUMBER OF INVERT FILE
RD	5000	N	8000	21	11

b．字段定义的基本情况（其中〝刊名〞字段为隐含检索字段 ）

FIELD NAME	LENGTH	REPEATABLE	NUMBER DATA TYPE	MANDATORY FIELD	INVERT
TITLE	250	N	N	Y	Y(ONL)
SUBTITLE	250	N	N	N	Y
PARALLEL TITLE	250	Y	N	N	Y
STATEMENTS OF RESPONSIBILITY	200	N	N	N	N
YEAR ISSUE	30	N	N	N	N
PLACE OF PUBLICATION	50	N	N	N	N
PUBLISHER	150	N	N	N	Y
COUNTRY	4	N	N	N	Y
YEAR OF PUBLICATION	20	N	N	N	N
LANGUAGE	4	N	N	N	Y
FREQUENCY	50	N	N	N	N
NOTE	250	N	N	N	N
ISSN	12	N	N	N	Y
CHINESE SERIAL NUMBER	16	N	N	N	Y
CLASSIFICATION	10	Y	N	N	Y(ONL)
SUBJECT HEADING	80	Y	N	N	Y(ONL)
LIBRARY COLLECTION	60	Y (SUBFIELD FIELD INDICATOR)			
LIBRARY CODE	6	N	N	N	Y
VOLUME ISSUE	20	N	N	N	N
YEAR	30	N	N	N	N
NUMERAL	2	N	Y	Y	N

过去《联合目录》的编辑工作完全用手工完成，是一次性工作，不具有数据的延续性。用计算机进行数据处理之后，不仅节省了人力和时间，而且使《联合目录》工作有了许多新的进展。主要体现在以下几点：a.增加了主题词和分类索引。在我们的医学期刊联合目录中，加这两项索引还是第一次，这是主观上想做，而手工难以实现的一件事。b．发挥 PRINT 模块的作用，输出格式严格向国际标准化著录格式看齐。例如：附件中的一些符号，·，:，/，=，·—，(，)，等，在录入时没有进行录入，是在 PRINT 模块中定义字段属性和对"文字"加以说明之后，由PRINT模块自动生成和输出的。c.由于使用 INDEX 模块对数据进行排序，彻底解决了手工排序错位的问题，同时提高了《联合目录》和图书馆的利用率。d．利用激光打印机输出，大大提高了制版质量。

2)建立数据库及数据处理过程

i）前处理工作：这项工作做得如何，直接影响着数据库的质量。工作做得细致、规范，数据的录入与维护就省时、顺利，数据的准确性就有保证，否则给校对和修改人员带来许多麻烦。本目录的前处理工作，主要由图书馆员完成。为记录标引主题词和分类号，这是一项技术性、专业性很强的工作，主要由副研究馆员完成。

ii）建立数据库：《联合目录》数据库与其它数据库的建库过程，在数据录入方面有着很大不同，主要反映在从微型机向小型机进行数据传送的过程上。软件人员首先在微型机上运用 dBASE III 建立临时数据库，供录入人员使用；在小型机上运用 MINISIS 建立正式数据库，接收批式输入（BATCHIN）的完整记录。采用这种方法的目的是：a．不受 HP3000 主机的直接控制，主机的开关与否，终端的个数多少，主机管理人员在否，均不影响录入人员在微型机上的工作。b．微型机上容易实现全屏录入（当时 MINISIS G 版还没有推出），比较非汉字终端，更显示出有中文提示的优越性。c.为避免录入工作的重复和错误，部分字段可以暂用代码代替，在通过中间程序(BASIC程序)产生 BATCHIN 格式文件的同时，将代码转为字段实际内容。

iii）数据的录入、转换及进一步修改：录入工作由图书馆"联合目录"专职人员完成。之后由软件人员编程，对数据进一步加工，输出 BATCHIN 的标准格式文件。最后将数据一次全部装入 MINISIS 数据库。近 3100 条记录，耗时 3 小时。数据的进一步修改，主要是对新增加的主题和分类两字段内容的分配进行系统地分析和检查。运用 INDEX 模块，将它们分别定为主键并拉平，以刊名为次键，全面、整体地检查主题词、分类号给定的是否准确合理。仅此一点充分反映和发挥了 MINISIS 软件的作用。

本数据库是在为 1991 年出版的 1986—1990 年的《联合目录》做准备，除馆藏字段没有输入数据外，其它字段内容已全部录入、修改完毕。待 1991 年将馆藏数据收集、整理、录入之后，即可出版新的《联合目录》。

《联合目录》数据库现已存在的数据，为本馆期刊自动化管理工作奠定了基础。

四。全国生物医学期刊资源分布调查数据库简介

本数据库与上述两数据库大同小异。它的收编范围是全国订购生物医学期刊的订户，对这些单位进行一年的调查统计。数据由中国图书进出口公司提供，这是一次性的资源分布调查工作，不带有延续性。它的机处理方式与《联合目录》类似，即：从微型机到小型机。数据库共设有 14 个字段，3931 条记录，目前数据已全部录入、修改完毕，正在进行数据统计、输出工作。

结束语

图书馆自动化是以计算机应用作为重要标志的。目前我馆的这项工作正在积极开展之中。期刊的联合目录工作为本馆的期刊自动化管理系统打下了一定的基础。中国生物医学文献数据库正在创建之中。待设备扩充之后，我们的图书馆集成管理系统将逐步行成和完善。

一九八九年十月于北京

附件：

《联合目录》输出格式

Output Format of UCBP

```
ACM Transactions on Programming Languages
  and Systems . --v.1(1979). --New York,
  USA : Association for Computing
  Machinery, 1979
  eng. --Quarterly
  ISSN 0164-0925                  738B033

A.I.Ch.E. Journal : Chemical Enginneering
  Research and Development. --v.1(1955).
  --New York, USA : American Institute of
  Chemical Engineering, 1955
  eng. --Bimpnthly
  ISSN 0001-1541                  810B001

AIDS Research : Acquires Immunodeficiency
  Syndrome. --v.1(1983). --New York, USA
  : Mary Ann Liebert,Inc., 1983
  eng. --Bimonthly
  ISSN 0737-6006                  631B029

AJNR : American Journal of Neuroradiology
  /American Society of Neuroradiology.
  --v.5(1984). --Baltimore, USA :
  American Roentgen Ray Society, 1984
  eng. --Monthly. --Continues: American
  Journal of Neuroradiology
  ISSN 0195-6108                  635B008
```

DATABASE OF CHINESE MECHANICAL ENGINEERING ABSTRACTS (DB CMEA)

Yang Zhe-zheng, Dong Chun-yuan, and Pei Xiao-ming
The Scientech Information Centre

1. Information Source

DB CMEA is a Chinese character data base built on HP3000/68 computer using MINISIS. The journal of CHINESE ENGINEERING ABSTRACTS is the information source for DB CMEA.

First I will give a brief account of the monthly journal of CHINESE MECHANICAL ENGINEERING ABSTRACTS.

The monthly journal of CHINESE MECHANICAL ENGINEERING ABSTRACTS was created in 1966. Each year it provides about 10, 000 abstracts and a copy of the subject index. It concentrates on collecting and reporting the literature and documents of the machinery industry and others related industries in China. The abstracts are taken from articles in periodicals, journals, trade magazines, monographs, conference proceedings, patent literatures, etc. The statistics in Table 1 show that articles from periodicals, trade magazines and journals are the main source of information for the data base. At present 760 different periodicals are available for abstracting in the editorial board. Among these, 250 are the key periodicals that the editorial board works with.

CHINESE MECHANICAL ENGINEERING ABSTRACTS (CMEA) is a comprehensive periodical dedicated to the machinery industry. It covers mechanical and electrical products in areas of material, manufacturing processes, design, manufacturing, research and development, production management. etc. The literature reported on in the journal fall into the 31 major categories shown in Table 2.

Table 1 Classification of the literature reported on CHINESE MECHANICAL ENGINEERING ABSTRACTS 1988

type of literature	number of articles reported	ratio (%)
articles from periodicals, journals, trade magazines, etc.	9890	89. 06
subject documents	487	4. 77
conference proceedings	289	3. 55
patent literatures	289	2. 83
total	11, 008	100

Table 2:CMEA documents distribution in a classified order

area	number of articles	ratio (%)
power engineering	759	7. 34
mechanical drive and general machine parts and components	653	6. 39
mechanical engineering materials	701	6. 86
casting process and equipment	570	5. 58
welding, cutting and bonding	448	4. 38
forging process and equipment	392	3. 48
metal corrosion and surface protection	384	3. 76
metallurgical machinery and equipment	368	3. 60
metal cutting machine-tools	365	3. 57
instrumentation	359	3. 51
fluid machinery	358	3. 50
heat-treatment process and equipment	351	3. 44
metal working	326	3. 20
motor vehicle and tractor	319	3. 12
petroleum and chemical machinery	299	2. 93
electronic technology, automation and computer	238	2. 23
cutter, abrasive, grinding wheels, gauges and measuring instruments	229	2. 24
construction machinery	228	2. 23
hoisting and handling machinery	217	2. 12
tools, fixtures, dies, jigs	206	2. 02

2. System functions

DB CMEA is in fact a comprehensive computerized information processing system which performs the retrieval book editing and database on-line retrieval functions. Figure 1 shows a flowchart diagram of the system. The document data, entered according to the required format of the worksheet (shown in Figure 2), are stored in tapes or disks after errors are checked and corrected. They are then sorted according to classification and a new abstracting number is given. The data are processed with abstracting and indexing programs written in FORTRAN (see Figure 3). The abstracting journal and the annual index are printed on a laser typesetter. The same data are stored on tapes or disks, after being processed by the adequate program, are converted for entry through BATCHIN (see Figure 4).

DB CMEA has 17 fields. It is retrieved by classification number, title, author and descriptor. For the sake of convenience to the user, the main retrieval path, which is the descriptor, is divided into three fields. The same literature can be searched by way of descriptor code, Chinese character and pinyin. Therefore, the data base has six inverted fields.

We have worked on the communication between the HP3000/68 and PC-XT computers. The personal computer with the Chinese character system is used as the terminal for HP computer allowing the retrieval results to be displayed on the screen as well as printed out. The system is characterized by stability, easy operation and fast response time. Figure 5 is a print-out sample of the retrieval results.

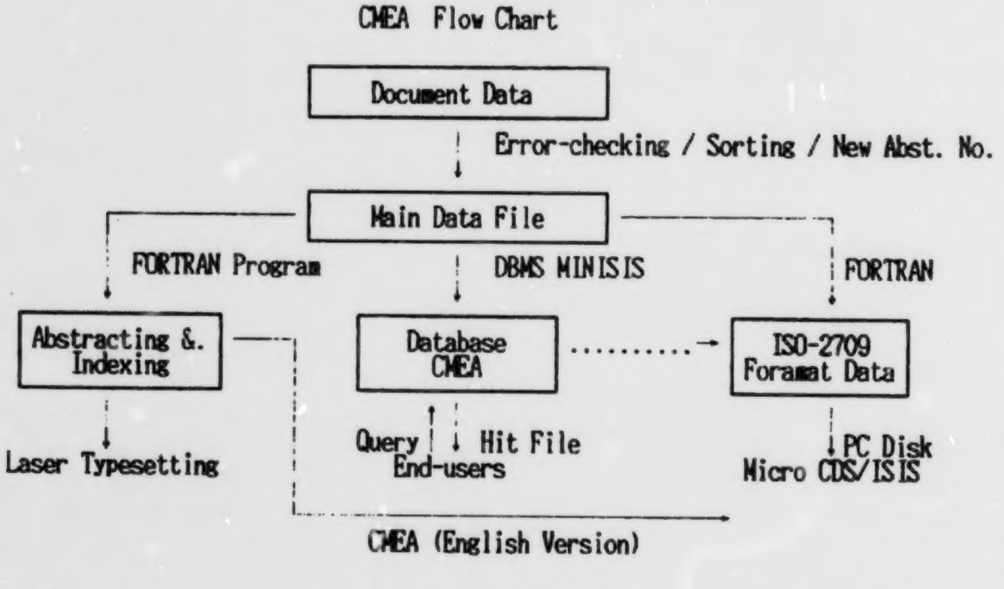
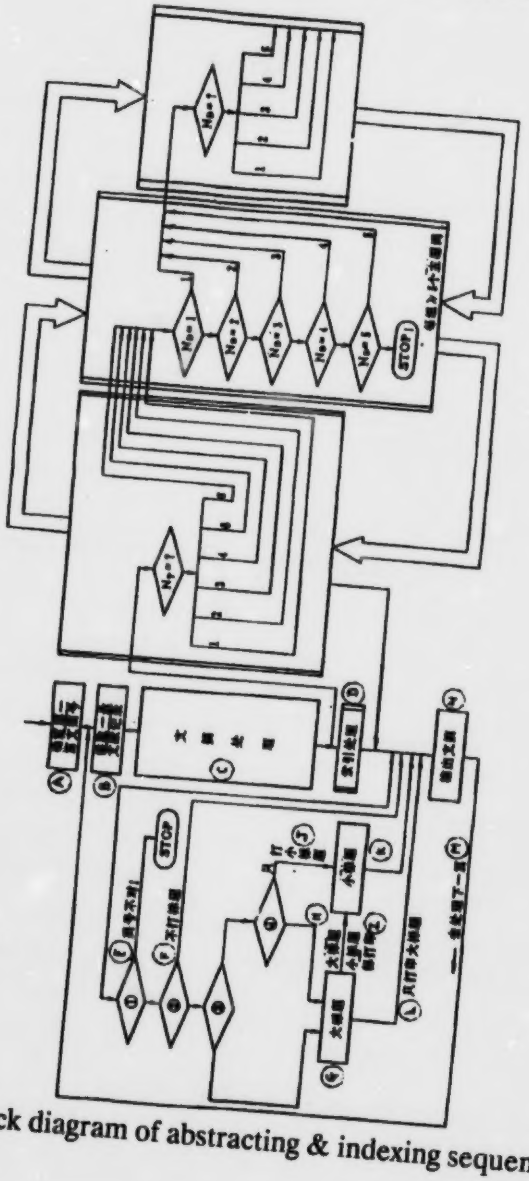

CMEA Flow Chart
Document Data
Error-checking / Sorting / New Abst. No.
Main Data File
FORTRAN Program
DBMS MINISIS
FORTRAN
Abstracting &. Indexing
Database CMEA
ISO-2709 Foramat Data
Laser Typesetting
Query
Hit File
End-users
PC Disk
Micro CDS/ISIS
CMEA (English Version)

Figure 1: System flowchart for DB CMEA

```
A 8811535
B 1721000
C 汽车离合器膜片弹簧的计算
D 刊
E 武鹏兰
K 北汽科技
L 北京汽车制造厂
P 26～31
X 图6
Y 表1
Z 参4
R 85-527-3/4
S 首先介绍了膜片弹簧设计计算的基本公式及参数选择方法，
  然后介绍了计算机程序并给出框图。最后研究根据计算结
  果如何确定膜片弹簧尺寸。
T129595÷ 24175÷ 26725÷ 32850
```

Figure 2: Data entries in the worksheet format



A. INPUT FIRST ABSTRACT NO.

B. ACCESS A RECORD

C:. PROCESS ABSTRACT

D. PROCESS INDEX

E. CLASSIFICATION NUMBER WRONG

F. NO-PRINT 1ST TITLE

G. PRINT 1ST TITLE (Y/N)

I. PRINT 1ST & 2ND TITLE

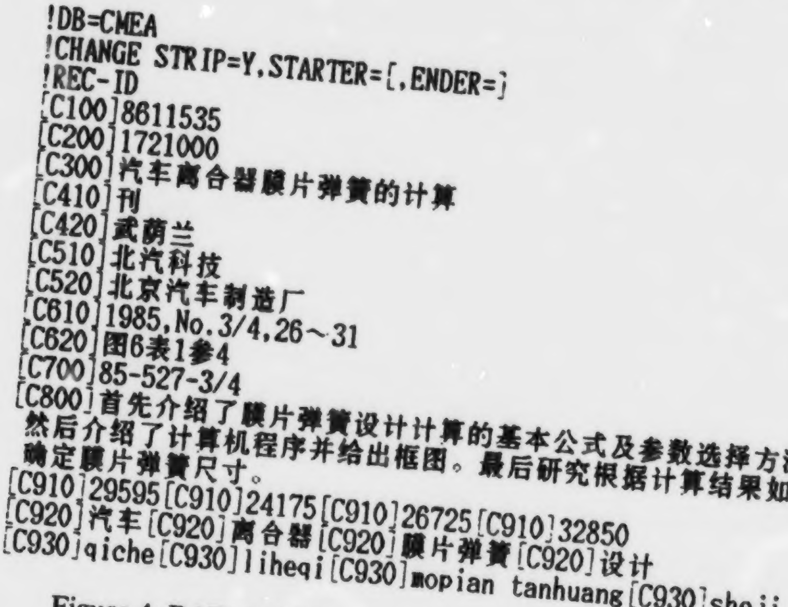
J. PRINT 2ND TITLE ONLY

L. PRINT 1ST TITLE ONLY

M. PROCESS NEXT ABSTRACT

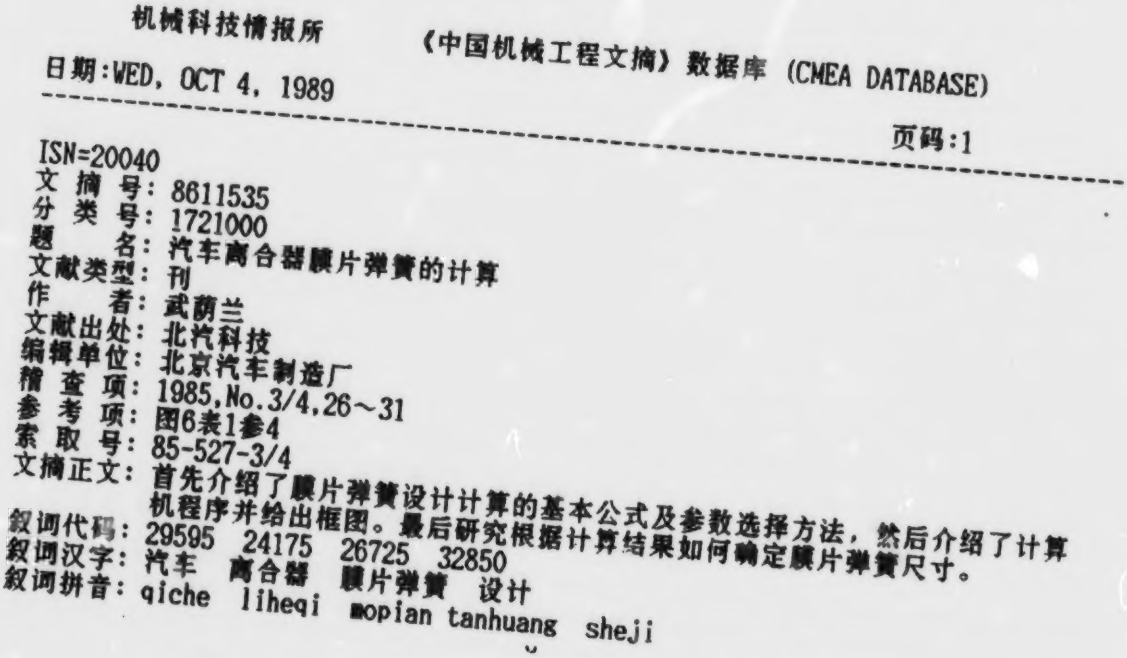
N. OUTPUT ABSTRACT TEXT

Figure 3: Block diagram of abstracting & indexing sequence



```
!DB=CMEA
!CHANGE STRIP=Y,STARTER=[,ENDER=]
!REC-ID
[C100]8611535
[C200]1721000
[C300]汽车离合器膜片弹簧的计算
[C410]刊
[C420]武荫兰
[C510]北汽科技
[C520]北京汽车制造厂
[C610]1985,No.3/4,26~31
[C620]图6表1参4
[C700]85-527-3/4
[C800]首先介绍了膜片弹簧设计计算的基本公式及参数选择方法，
然后介绍了计算机程序并给出框图。最后研究根据计算结果如何
确定膜片弹簧尺寸。
[C910]29595[C910]24175[C910]26725[C910]32850
[C920]汽车[C920]离合器[C920]膜片弹簧[C920]设计
[C930]qiche[C930]liheqi[C930]mopian tanhuang[C930]sheji
```

Figure 4: BATCHIN format data for entry into DB CMEA



```
    机械科技情报所          《中国机械工程文摘》数据库 (CMEA DATABASE)
日期:WED, OCT 4, 1989
-------------------------------------------------------------------------
                                                              页码:1
ISN=20040
文 摘 号: 8611535
分 类 号: 1721000
题    名: 汽车离合器膜片弹簧的计算
文献类型: 刊
作    者: 武荫兰
文献出处: 北汽科技
编辑单位: 北京汽车制造厂
稽 查 项: 1985,No.3/4,26~31
参 考 项: 图6表1参4
索 取 号: 85-527-3/4
文摘正文: 首先介绍了膜片弹簧设计计算的基本公式及参数选择方法，然后介绍了计算
          机程序并给出框图。最后研究根据计算结果如何确定膜片弹簧尺寸。
叙词代码: 29595  24175  26725  32850
叙词汉字: 汽车  离合器  膜片弹簧  设计
叙词拼音: qiche  liheqi  mopian tanhuang  sheji
```

Figure 5: Print-out sample of retrieval results

3. The development and application of the system

3. 1. The publication of CHINESE MECHANICAL ENGINEEERING ABSTRACTS of English edition

Based upon the Chinese edition of CHINESE MECHANICAL ENGINEERING ABSTRACTS, in 1987 we started to publish the English edition using this system. The English edition is a quarterly journal. The worldwide distributor is Pergoman Inc. The tapes in 150/2709 format can be provided to the user.

3. 2. Information transfer by floppy disk

The data from DB CMEA are retrieved and down-loaded onto floppy disk in the standard format according to the requirement of the user in a specific fields of interest. The disk is then provided to the user so as to facilitate the information collection and creation of data bases on personal computers for organizations that do not have access to the necessary information source. Thus, information can be shared.

3. 3. Computer-aided documentation statistics and analysis

The editing of the Chinese and English editions of CHINESE MECHANICAL ENGINEERING ABSTRACTS, the retrieval from Chinese character data bases, information down-loading services are all related to the application of the system. In addition, the editorial board uses the system to perform some of the information statistics and analysis services.

In the past, the retrieval tools in book form used to be edited manually. With the publication of the index for the CHINESE MECHANICAL ENGINEERING ABSTRACTS, the editorial board had to do a certain statistical and analytical work on the data reported. Because of large amount of the work involved, only part of statistics and analysis could be performed. Now with help of the computer the statistics and analysis can be performed in a systematic way and with more accuracy.

We have noticed that in recent years, considerable amount of statistics and analysis have been performed with the help of computers and data bases. Examples are the analysis of the R & D program on tribology performed by the Material Research Institute of Federal Republic of Germany with the help of subject headings(1), the analysis of Japanese R & D activities by Tokyo University using INSPECT and CAC data bases (2), research on firm and industry opportunities influencing publications performed by Pennsylvania University of the United States using SCISEARCH and BRS/PATSEARCH(3), the research on the increase of patent on high polymer using the HEMYA (CHEMISTRY) data base (4) performed by the Soviet Union, the statistics of the publications by Polish scholars in foreign countries and their distribution using INSPECT, CAC AND SCI (5).

The computer-aided document statistics and analysis is not a simple document retrieval. It is a combination of the retrieval and analysis of documents. It is a pool of efforts by Scientists and documentalists. It is an expansion and extension of the retrieving functions of the data bases. The role played by the data base in technical and scientific development can be explored in a much more significant way.

We are of the opinion that the statistical analysis we so far performed on the CHINESE MECHANICAL ENGINEERING ABSTRACTS is of some value and significance to the exploration of rules for document indexing, the use and revision of thesaurii and understanding certain trends of development in the machinery industry.

1. The determination of the optimal coordinate elements in indexing by key words

The document subject is always very complicated and there are seldom any cases where the subject can be expressed by only one keyword. It does not necessarily mean that the more keywords the better the results. It is really a very practical question as to how many keywords should be used in the preparation of the retrieving tools. For years, the journal of CHINESE MECHANICAL ENGINEERING ABSTRACTS used a three coordinate elements keyword system. But since 1983 when the computer was introduced for the preparation of the indexing, the five coordinate elements system was adopted instead. Here is an example of the five coordinate element keyword system:

BORIDING
brittleness/laser heating/heat treatment/case depth 8307405

Where boriding is the primary searching term and "brittleness", "laser heating", "heat treatment" and "case depth" are secondary searching terms.

We have statistics on the coordinate keyword elements used in the 30,000 titles of the CMEA documents over 1983, 1985 and 1986 (see Table 3). We found that only 6.2% of the documents used one keyword for indexing. The largest number, 79.4%, used two to three keywords. About 11. 3% of the documents used four keyword element for the indexing. A few of the documents used five keyword element for the indexing. The same can be seen from the curve about the coordinate keyword element system (see Figure 8). When more than three coordinate keyword elements are used, the curve begins to flatten out. The use of multiple coordinate keywords that are of more than four does not help much to improve the coverage rate. The use of three coordinate keyword elements can help cover 85. 6% of the document subjects. Our conclusion is the use of the three coordinate keyword elements is the practical and optimal option. The number of the keywords can certainly be increased somewhat if the computer is used for the indexing. However, the use of more than four keywords does not significantly improve the coverage of the document subjects.

Table 3. The coordinated relationship of the number of the keywords for indexing the document subjects

year	1983		1985		1986		total	
number of keywords	indexing entry	%	indexing entry	%	indexing entry	%	indexing entry	%
1	1495	7	1576	7. 1	1035	4. 6	4106	6. 2
2	8485	41	8936	40	6960	31	24381	37. 3
3	8090	38	11166	50	8572	38. 2	27828	42. 1
4	2392	11	582	2. 6	4482	20	7456	11. 3
5	431	3	71	0. 3	1370	6. 1	1872	3. 1
total	20893	100	22331	100	21419	100	65643	100

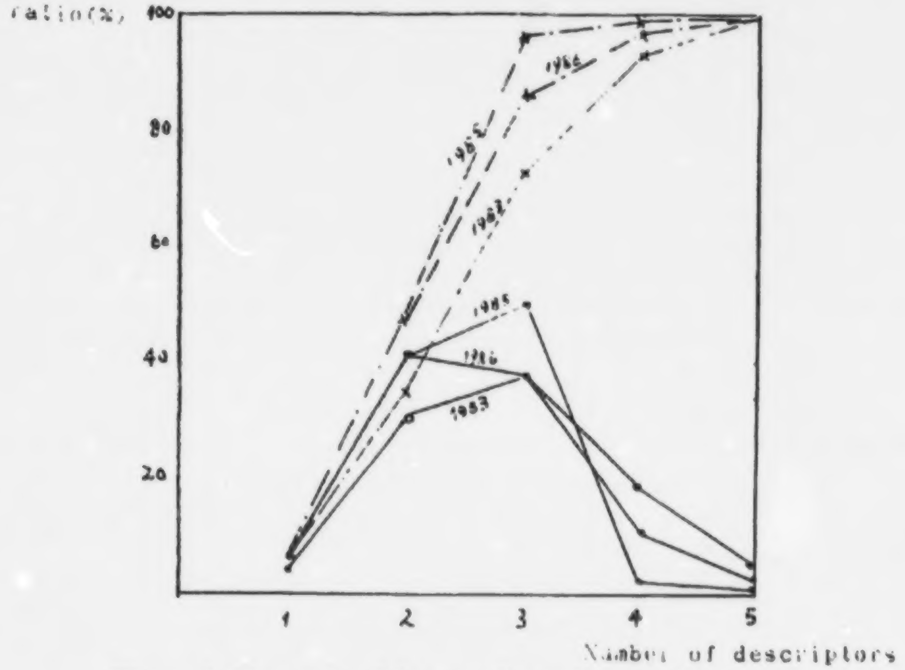


Figure 6: The curve of the coordinate keyword elements

Note: The solid line represents the percentage of different coordinate keyword elements in the CMEA annual indexer and the dashed line represents the coverage of the document subjects. The data for the curve are taken from the statistics of 1983, 1985, and 1986.

2. The revision of the subject headings

The statistics of the keyword usage frequency in CMEA is not only an important basis for the addition and deletion of relevant keywords, it is also a very valuable reference for the revision of the subject headings. We found that some of the keywords have a very high frequency of usage. The keyword "energy conservation" appeared 217 times (used for checking) in the 12 issues of the CMEA abstracts. The keyword "nodular cast iron" appeared 140 times. This shows that energy conservation is an important issue for various departments of the machinery industry and the nodular cast iron is a very widely used material. But, on the other hand, we should also see that the reason why some of the keywords are very frequently used is because the selection and creation of the keyword is not detailed enough. In fact, energy conservation is a general concept. Water conservation, electrical power conservation, oil conservation, coal conservation and many others all fall into the category of the energy conservation. Similarly, nodular cast iron can be further divided into alloy nodular cast iron, rare earth metal nodular cast iron and others. This shows that the detailed classification principles have to be worked out for the subscript and superscript keywords for those widely used technical concepts and terminology so as to help improve the document searching accuracy.

3. Understanding of certain technical development trends for the machinery industry

Figure 7 shows the keywords that have a usage frequency of more than 50 (1983's statistics). The high technology keywords such as "finite element analysis", "computer aided design" have a very high frequency of usage. But majority of the terms that we deal with daily are related to the traditional manufacturing processes (like heat treatment, welding, foundry, grinding, etc.), materials (like nodular cast iron, stainless steel, steel, aluminum, etc.), key parts and components (like gears, crank shafts, bearings, blades, etc.) and major products (like pressure vessels, asynchronous motors, steam turbine generators, motor vehicles, pressure vessels, etc.). Although what is listed in Figure 7 is the statistics of a whole year, it does tells us that these manufacturing processes, materials, parts and products are required in large quantities and closely related to many branches of the national economy. They really have very important significance to the machinery industry. It is erroneous to simply develop the high technology at the expense of the traditional technology and it is also wrong if we neglect and underestimate the development of high technology and the role played by the high technology in the modernization of the traditional technology. These statistics provides a basis and guide for the formulation of the policy of science and technology.

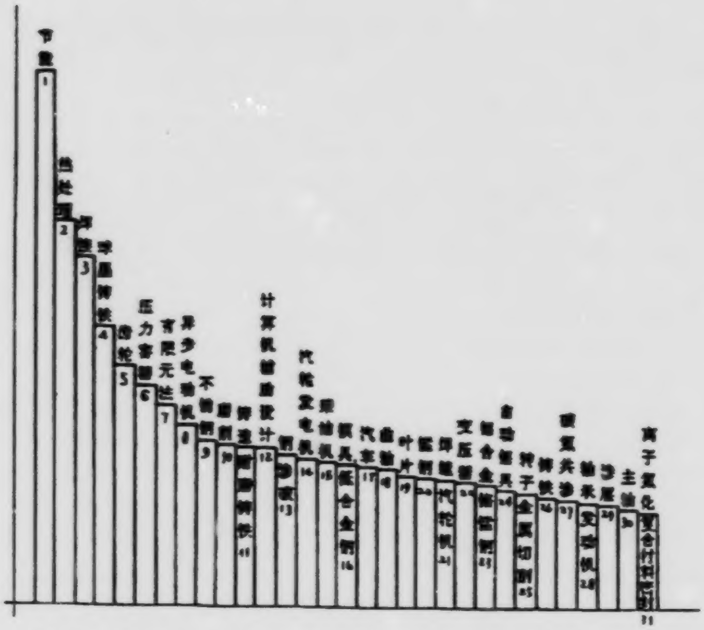


Figure 7: Keywords with a frequency greater than 50

1. Energy conservation
2. Heat treatment
3. Welding
4. Nodular cast iron
5. Gear
6. Pressure vessel
7. Finite elements analysis
8. Asynchronous motor
9. Stainless steel
10. Grinding
11. Foundry/wear-resistant casting
12. CAD
13. Steel carbonization
14. Steam turbine generator
15. Diesel engine
16. Die and molds
17. Motor vehicle
18. Crank shaft
19. Blade
20. Manganese steel
21. Welding joint/steam turbine
22. Transformer
23. Aluminum alloy/Cr-Mn steel
24. Automatic drilling tools
25. Rotor/metal machining
26. Castings
27. Nitrocarburizing
28. Bearing/engine
29. Case depth
30. Main spindle
31. Ion nitriding/composite material searling

REFERENCES

1. Tribologie(Reibung, Verschleiss, Schmierung), Band 1, 1981 P. 39-60
2. Translation of Science and Technology(in Chinese) , 1985, N0. 2
3 R & D Management 17, 3, 1987, P167-173
4. HTN Cepl, 1983, No. 8
5. HTN, Cepl, 1983, No. 112

《中国机械工程文摘》数据库

杨则正、董春元、裴晓明

1、信息源

《中国机械工程文摘》数据库，是利用 MINISIS 软件在 HP3000／68 计算机上建立的中文数据库．这个数据库以《中国机械工程文摘》杂志为信息源．这里，首先对《中国机械工程文摘》杂志作一介绍．

《中国机械工程文摘》杂志创办于 1966 年，月刊，年文献报导量约 10000 篇，每年相应出版主题索引一册，重点收集、报导我国机械和与之密切相关的文献，取材于期刊论文，专题文献，会议论文，以及专利文献．统计表明，期刊论文是最主要的信息源（表 1），目前，供编辑部选用的中文期刊有 760 种，其中 250 种是核心期刊．

《中国机械工程文摘》是机械类综合性杂志，内容涉及机电产品的材料工艺、设计、制造、科研和生产管理等各个领域，报导文献包括 31 个大类．表 2 是《中国机械工程文摘》报导文献的专业分布情况．

2、系统的功能

《中国机械工程文摘》数据库，实际上是计算机支持的综合性的文献处理系统．系统具有编辑书本式检索工具的功能和数据库联机检索的功能．图 1 是系统工作流程．按工作单格式录人的文献数据（图 2），经查错和改错，按类号排序，并赋于新文摘号后，存储于磁带（或磁盘），用 FORTRAN 编写的文摘和索引编辑程序对这些文献数据进行处理，（图 3）再经激光照排处理，就完成文摘刊物和年度索引的制作．同样是记录在磁带（或磁盘）上的这些文献数据，经过相应程序的处理，便得到 MINISIS 接受的 BATCHIN 格式的人库数据（图 4）．

《中国机械工程文摘》数据库共设 17 个字段．通过分类号，题名，作者和主题词检索．考虑到用户检索习惯和操作方便起见，又将主题词字段（主要检索途径）分成三个字段，用户可以从主题词号，汉字和拼音不同途径查找到同一篇文献．因此，实际上，数据库有 6 个字段是倒排字段．

在硬件方面，我们采取了一些措施，实现了 HP 机与 PC／XT 微机的通讯．由于利用配备有汉字系统的微机作为 HP 机的终端，用户检索结果既可以显示在屏幕上，也可以通过微机配置的打印机打印出来．系统稳定，操作方便，响应时间令人满意．图 5 是检索结果打印示样．

表 1　《中国机械工程文摘》文献类型
(以 1988 年为例)

文献类型	篇数	比例 (%)
期刊(包括学报)论文	9890	89.06
专题文献	487	4.77
会议论文	342	3.55
专利文献	289	2.83
总计	11008	100.00

表 2　《中国机械工程文摘》文献按专业分布
(1988 年为例)

专业范围	篇数	比例 %
动力工程	759	7.34
机械工程材料	701	6.86
机械传动与通用零部件	653	6.39
铸造工艺与设备	570	5.58
焊接、切割与贴接	448	4.38
铸造工艺与设备	392	3.84
金属腐蚀与表面保护	384	3.76
冶金机械	368	3.60
金属切削机床	365	3.57
仪器仪表	359	3.51
流体机械	358	3.50
热处理工艺与设备	351	3.44
金属切削加工	326	3.20
汽车、拖拉机	319	3.12
石油化工机械	299	2.93
电子技术、自动化、计算机	238	2.33
刀具、磨料、磨具、量具和量仪	229	2.24
建筑工程机械	228	2.23
起重运输机械	217	2.12
工具、夹具、模具	206	2.02
合计	11008	100.00

图 1　中国机械工程文摘计算机处理系统流程图

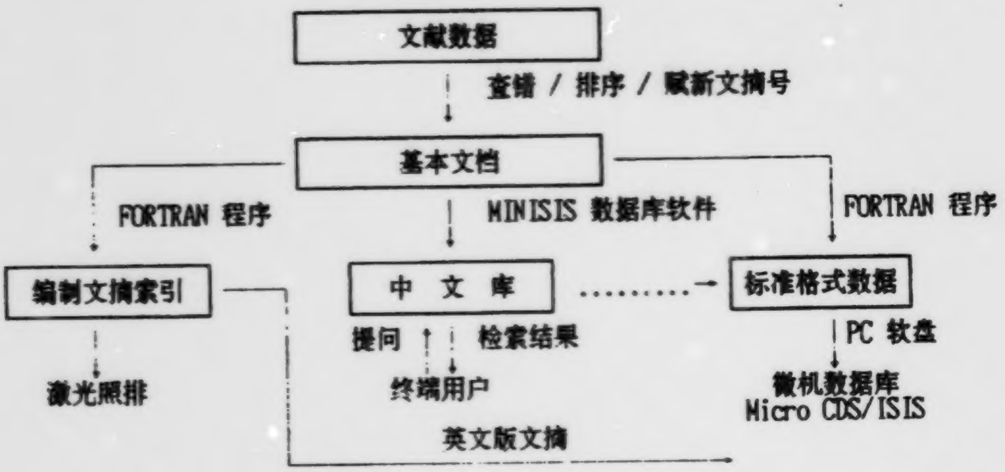
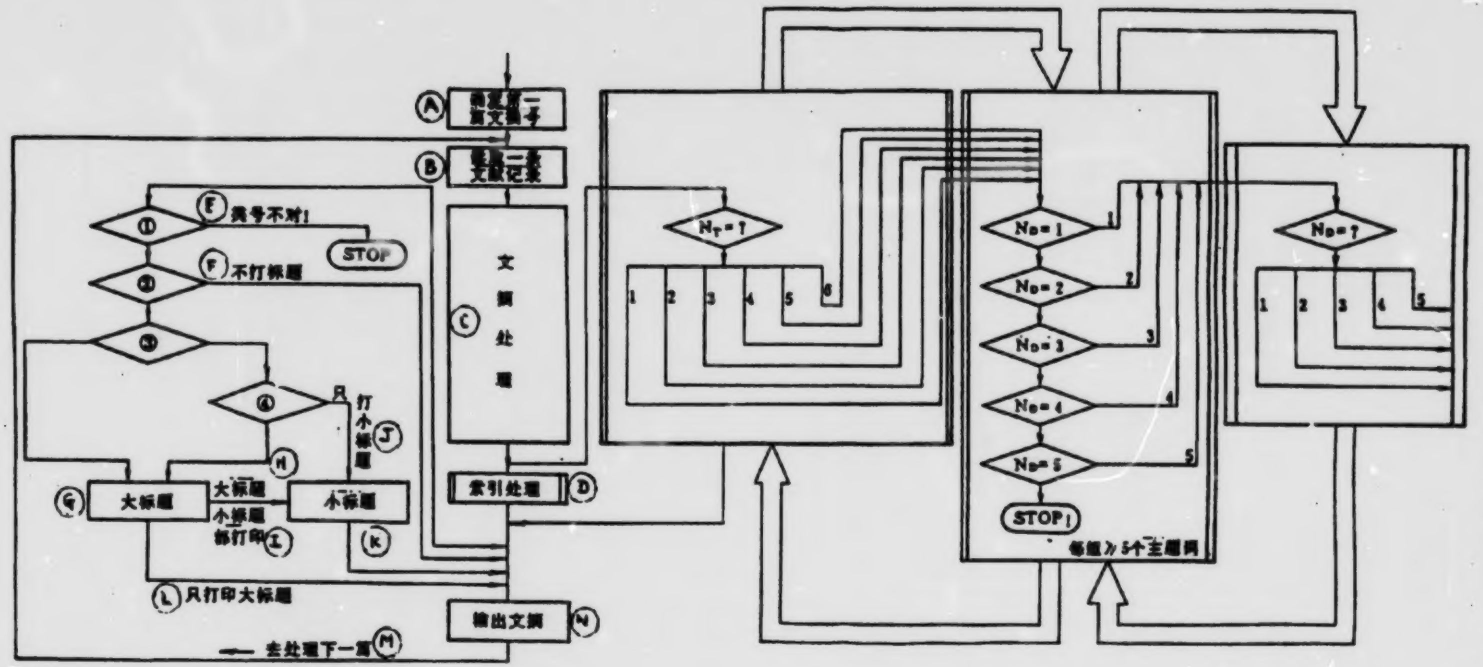

文献数据
查错 / 排序 / 赋新文摘号
基本文档
FORTRAN 程序
MINISIS 数据库软件
FORTRAN 程序
编制文摘索引
中　文　库
标准格式数据
提问
检索结果
终端用户
激光照排
PC 软盘
微机数据库
Micro CDS/ISIS
英文版文摘

图 2　按工作单格式录入的文献数据

A 8611535
B 1721000
C 汽车离合器膜片弹簧的计算
D 刊
E 武荫兰
K 北汽科技
L 北京汽车制造厂
P 26～31
X 图6
Y 表1
Z 参4
R 85-527-3/4
S 首先介绍了膜片弹簧设计计算的基本公式及参数选择方法，然后介绍了计算机程序并给出框图。最后研究根据计算结果如何确定膜片弹簧尺寸。
T129595÷ 24175÷ 26725÷ 32850

图 3　文摘和索引编排程序框图


Ⓐ
Ⓑ
文摘处理 Ⓒ
索引处理 Ⓓ
输出文摘 Ⓝ
①
Ⓔ 类号不对!
STOP
②
Ⓕ 不打标题
③
④
只
打小标题 Ⓙ
Ⓗ
Ⓖ 大标题
大标题
小标题
都打印 Ⓘ
小标题 Ⓚ
Ⓛ 只打印大标题
去处理下一篇 Ⓜ
N_T = ?
1
2
3
4
5
6
N_D = 1
N_D = 2
N_D = 3
N_D = 4
N_D = 5
STOP!
每组≯5个主题词
N_D = ?
1
2
3
4
5

① 类号在表内？　② 等于上一类号？　③ 大小标题打印判断？

图 4　MINISIS 接受的 BATCHIN 格式的入库数据

```
!DB=CMEA
!CHANGE STRIP=Y,STARTER=[,ENDER=]
!REC-ID
[C100]8611535
[C200]1721000
[C300]汽车离合器膜片弹簧的计算
[C410]刊
[C420]武荫兰
[C510]北汽科技
[C520]北京汽车制造厂
[C610]1985,No.3/4,26~31
[C620]图6表1参4
[C700]85-527-3/4
[C800]首先介绍了膜片弹簧设计计算的基本公式及参数选择方法,
 然后介绍了计算机程序并给出框图。最后研究根据计算结果如何
 确定膜片弹簧尺寸。
[C910]29595[C910]24175[C910]26725[C910]32850
[C920]汽车[C920]离合器[C920]膜片弹簧[C920]设计
[C930]qiche[C930]liheqi[C930]mopian tanhuang[C930]sheji
```

图 5　数据库检索结果打印式样

```
      机械科技情报所         《中国机械工程文摘》数据库 (CMEA DATABASE)

日期:WED, OCT 4, 1989                                         页码:1
------------------------------------------------------------------------

ISN=20040
文 摘 号: 8611535
分 类 号: 1721000
题    名: 汽车离合器膜片弹簧的计算
文献类型: 刊
作    者: 武荫兰
文献出处: 北汽科技
编辑单位: 北京汽车制造厂
稽 查 项: 1985,No.3/4,26~31
参 考 项: 图6表1参4
索 取 号: 85-527-3/4
文摘正文: 首先介绍了膜片弹簧设计计算的基本公式及参数选择方法，然后介绍了计算
          机程序并给出框图。最后研究根据计算结果如何确定膜片弹簧尺寸。
叙词代码: 29595  24175  26725  32850
叙词汉字: 汽车  离合器  膜片弹簧  设计
叙词拼音: qiche  liheqi  mopian tanhuang  sheji
```

3　系统的发展和应用

3、1　出版英文版《中国机械工程文摘》

在中文版《中国机械工程文摘》的基础上，1987 年开始，利用计算机编辑英文版《中国机械工程文摘》（季刊），由英国培格曼公司向国外销售，并可向用户提供 ISO／2709 格式的文献磁带．

3、2　开展文献套录服务

根据用户需要，将《中国机械工程文摘》特定专业领域（例如仪表）的文献检索出来，构造成标准格式记录在软盘上，向用户提供，减轻基层文献机构信息采集和建立专业微机数据库的工作量，实行资源共享．

3、3　计算机辅助的文献统计分析

前面讲到的编辑中英文版《中国机械工程工程文摘》杂志，中文库检索，文献套录，都是系统的应用．除此之外，编辑部本身也利用系统进行一些文献统计分析工作．

以前，我们采用手工方式编辑书本式检索工具．那时候，伴随着《中国机械工程文摘》年度索引的出版，编辑部每年通常都要做一些文献统计分析．由于手工工作量太大，只限于部分的分析．现在，借助于计算机的帮助，有可能进行较为系统的准确的统计分析．

我们注意到，利用计算机和数据库进行文献统计分析的工作，近年来在国外得到普遍的开展．例如，联邦德国材料试验研究所借助主题词分析摩擦学研究促进计划 (1)；日本东京大学利用 INSPEC 和 CAC 数据库，分析日本科技人员的"活动能力" (2)，美国宾夕凡尼亚大学利用 SCISEARCH 和 BRS／PATSEARCH，研究公司和行业特点与科技出版物的关系，苏联利用《ХИМИЯ》数据库，研究高分子聚合物专利增长情况 (4)，波兰利用 INSPEC，CAC，SCI，研究波兰学者在国外发表的科技文献数量和分布情况 (5)．

计算机辅助的文献统计分析，不是单纯的检索，而是文献检索和统计分析结合，科技人员与文献工作考结合，是数据库检索功能的扩大和延伸，使数据库在科学技术发展中的作用得到更充分的发挥．

我们认为，我们对《中国机械工程文摘》进行的统计分析，对于探讨文献标引工作的规律，主题词表的使用和修订，了解机械工业发展的某些动向，都有参考价值和意义．

1、主题法标引中最佳的组配元数的确定

文献主题比较复杂，一个主题词就能表达主题的文献不多，但是，也不是主题词越多越好．检索工具究竟采用几元组配，在编制检索工具的工作中，这是一个很实际的问题．多年来，《中国机械工程文摘》采用三元组配，自 1983 年开始利用计算机编制索引以后，采用五元组配．下面是使用五个主题词的标引示例:

渗硼

脆性 / 激光加热 / 热处理 / 渗层　　8307405

这里，“渗硼”是检索词，“脆性”，“激光加热”，“热处理”和“渗层”是说明词.

我们对《中国机械工程文摘》1983，1985，1986 三年报导的 30000 余篇文献的主题词组配关系统进行了统计（表 3），发现只用一个主题词标引的是少数（6.2%），使用二个或三个主题词组配最多，占 79.4%，采用四个主题词也不多（11.3%），五个主题词的极少（只占 3.1%），从组配关系曲线（图 6）也可以看出，超过三元组配，曲线已逐渐平坦．采用四元以上的多元组配对提高复盖率的作用不大．采用三元组配，可使 85.6%的文献的主题能表达出来．结论是，采用三元组配，是可行的最佳方案，采用计算机编制索引，虽然主题词数目可适当增多，但超过 4 个主题词，对提高文献复盖率的意义不大．

2、词表修订与主题词增删

《中国机械工程文摘》主题词使用频率统计，不仅成了《机械工程主题词表》主题词增删的重要依据，而且为词表修订提供了有价值的材料．例如，我们发现，有些主题词使用频率很高．“节能”在 1983 年 12 期文摘中出现 217 次（作检索词），“球墨铸铁”出现 140 次．一方面说明节能是机械工业各部门关注的问题，球墨铸铁应用广泛．另一方面，应看到造成主题词频率使用很高的原因还在于，词表中有些主题词选词和立词不细．实际上，节能是泛指概念，节电，节水，节油，节煤都属于节能的范畴．同样，球墨铸铁有合金球墨铸铁、稀土铸铁等等．这说明，应对节能，球墨铸铁之类使用广泛的技术概念和术语，确定上位词和下位词的细分原则，以提高文献类查准率．

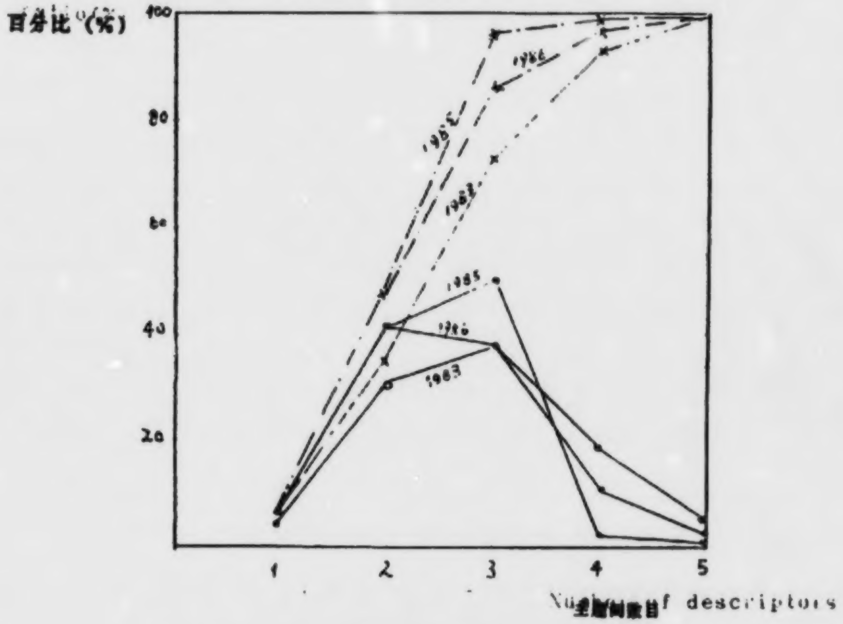
3、4　了解机械技术发展的某些动向

图 7 是使用频率在 50 次以上的主题词（1983 年数字）．象“有限元素”和“计算机辅助设计”这样高技术的主题词使用频率很高．但是，绝大部分是传统工艺（例如，热处理，焊接，铸造，磨削等），材料（例如球墨铸铁，不锈钢，钢材，铝合金等），关键零部件（例如，齿轮，曲轴，轴承，叶片等），以及主要产品（例如，压力容器，异步电动机，汽轮发电机，汽车，变压器等）．尽管只是一年的统计数字，但说明，这些大量面积的现有工艺、材料、零部件和产品，对机械工业有重大意义，撇开传统技术一昧发展高技术是错误的，同样，无视高技术的发展，看不到高技术的发展以及高技术在改造传统技术中的作用，也是错误的．事实上，这些统计数字，为制订科学技术政策提供了依据．

表 3　不同主题词数目的组配关系

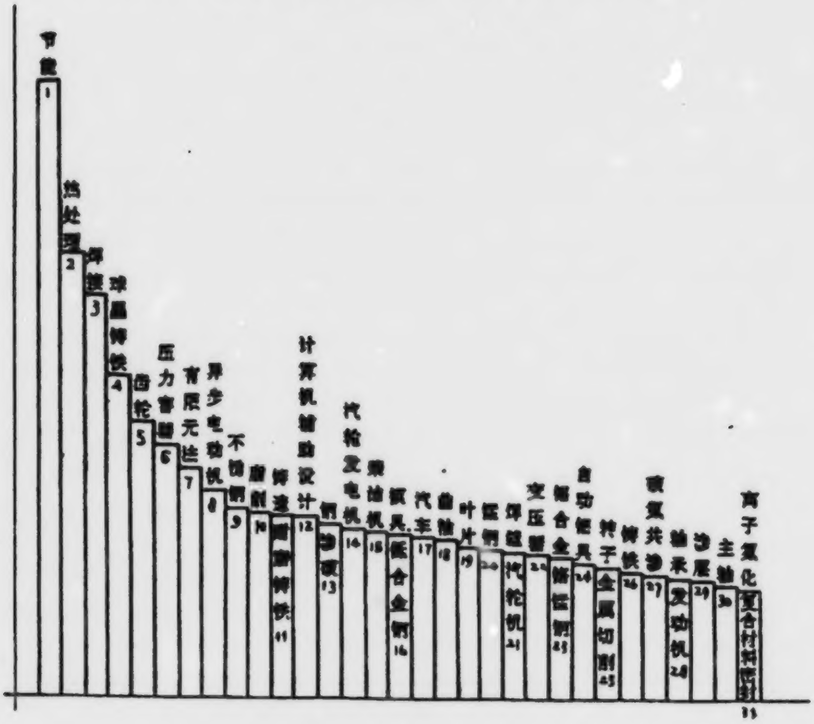
年份	1983		1985		1986		总　计	
主题词数目	索引款目数	%	索引款目数	%	索引款目数	%	索引款目数	%
1	1495	7	1576	7.1	1035	4.6	4106	6.2
2	8485	41	8930	40	6960	31	24381	37.3
3	8092	38	11166	50	8572	38.2	27828	42.1
4	2392	11	582	2.6	4482	20	7456	11.3
5	431	3	71	0.3	1370	6.1	1872	3.1
合计	20593	100	22331	100	21419	100	65643	100

图 6　主题词组配关系曲线



说明：实线表示不同主题词数目组配的索引项目在《年度索引》中的百分比。点划线表示反映文献主题的覆盖率。图中反映 1983、1985、1986 三年统计结果。

图 7 使用频率大于 50 次的主题词



参 考 文 献

1. TRIBOLOGIE (Reibung. Verschleiss. Schmierung) .Band 1 1981, S.39～60
2.科学学译丛 1985, No2
3. R&D Mangagement 17, 3, 1987, P167～173
4. HTИ Cep. 1, 1983, No.8
5. HTИ, Cep. 1, 1983, No11.

Building Domestic and Foreign Instrument Products Data Base through use of Minisis

Wu Xiaohong and Sun MengLan
Instrumentation General Technology and Economics Institute Ministry of Machinery and Electronics Industry

[Abstract] This article describes the procedure of setting up Domestic and Foreign Instrument Products Data Base, including data organization, data processing, and the function of building KSAM data base.

Introduction

Instrument industry is a knowledge-intensive and technology-intensive industry which involves a wide range of disciplines and technologies. With the progress of new technologies, the instruments have played an increasingly important role in the fields of new and developing technologies in recent years, and have found usage in various sectors of national economy.

The instrument products feature a great many of categories and varieties, and quick replacements of older generations by new ones. In order to enable the instrument users and the production, scientific reasearch, and management departments to search out the needed technical and economic information of the products from a great many of product informations so as to meet the needs of type selection, R D, and management, and to facilitate macro-management and micro-information guidance, we have set up the Domestic and Foreign Instrument Products Data Base.

In order to meet the needs of the said departments, we made a systematic analysis on the basis of our investigation. We have collected the information about the product manucturers, and about 40 data items of products, such as usage, main technical features, technical indicators, working principle, construction features, new technology, completeness, reliability, working condition required, serialization, grade of standard, development form, technical level, sales volume, price, and cost. We set up the Data Base through use of MINISIS and the data are stored according to variety or specifications.

I. Data organization and processing

The instrument products are divided into 13 categories, 146 sub-categories, 776 series, 17,000 varieties, 100,000 specifications. The criteria of different categories divided into sub-categories may differ. Some are divided according to the function or completeness of the system of variables measured, such as industrial automation instruments; some are divided according to the usage, methods of use, and application scope of the products, such as electrical instruments; and some are divided according to the construction principle and application fields of the products, such as optical instruments. Similarly, the criteria of sub-categories divided into series, and series into varieties are also different.

As the instruments involve many diciplines and technologies, the functions of the products and the physical quantities in the technical parameters involve electromagnetics, optics, acoustics, heat, bionics, chemistry, oceanography, geophysics, cybernetics, system engineering theory, etc.

Facing such complicated and non-document data that have to meet the different needs of the users, how are we to organize and process the data so as to meet the needs of both macro-statistics and micro-searching?

One code for one item is the best way, but as there is no such coding in the GB and JB while the product coding in the national industrial and agricultural products (commodities and materials) classification and code (GB7635-87) used domestically has only attained to the level of series. According to the statistics of 1985, the number of varieties under the series is only up to 6000, and the number of specifications is only up to 190,000. It is obvious that only using product coding will not do. On the basis of the systematic analysis and according to the features of the instrument products, we adopt the form of combining classifaction with descriptor method, series classification and variety use code form, while in the method of descriptor the key words and technical data are extracted from the related fields by making use of the word-eztracting function of Minisis. So macro-searching can be carried out by using classification while micro-searching can be carried out by using the descriptor method.

In the main data base, there are three fields that need to be extracted and inverted, and their contents are different, including the usage of the products, main technical preformance and technical indicators, key technology and new technology.

The extract of the two fields the usage of products and key technology and new technology is equal to free word indexing. The extract of the main technical performance and technical indicators is a complicated task, because the technical parameters of the instruments involve physical quantities of various disciplines, the contents are complicated, and the methods of extract are different, which will directly influence the size of inverted file and search speed. After analyzing a vast amount of data, we put forward a feasible extract method for identifier, parameter selection, and standardization, etc. It has proved quite effective in practical use and will be further perfected. Among the other fields, there are three that have adopted the GB code and standardized writting-method . One has adopted the code of Information Bureau of MMEI of PRC; other one has adopted the writing method stipulated by the State Science and Technology Commission; and two have defined their own standardized writing methods.

II. The Function of KSAM

There are name, address, telephone number, telex cable and other related informations of each enterprise in the main database. There are many varieties and specifications of products for one enterprise. In order to avoid inputing the same data repeatedly, and to minimize the errors caused by the repeated input, we have built a data base of basic conditions of enterprises (KSAM) taking enterprise's name-code as the main key word, and taking the KSAM data base as validation file of enterprises' name-code in the main data base. We only input the enterprises name-code to the main base, because the other information can be obtained from the KSAM.

It appears more important to build KSAM for the input of overseas products' information. We collected the products of some dozen countries, such as America, U.K., Japan, France, U.S.S.R., W. Germany and E. Germany, which brought about many languages, so the error rate of input became higher. Since building the KSAM, it has not only solved this problem, but also the problem of mnay names for one enterprise.

The building of KSAM has also improved working efficiency by avoiding repeated data input.

III. Conclusion

The instrument products data base uses the retrieval method of inversion to extraction of key words, filling up the void of micro-searching to the available goods and meterials' code formalated by the government for the computer retrieval. To ensure the source, channel base, flow, we shall organize the correspondent information network, and renew the data and inquire of the outside world from time to time.

The instrument products data base, as one of the basic data bases of the technical and economic information system of the instrument industry, will be developed continuously.

利用MINISIS建立仪器仪表国内外产品数据库

吴晓红、孙梦兰

[文摘]:本文简述了利用MINISIS的功能建立仪器仪表产品数据库的情况，从数据组织和处理、KSAM库的建立等方面进行了介绍。

一、概述

仪器仪表工业是知识密集、技术密集产业，涉及到多种学科和技术，随着世界新技术的进步，仪器仪表近几年在新兴技术领域中的作用越来越重要，其用途涉及到国民经济的各个部门。

仪器仪表产品的特点是品种门类繁多，产品更新换代快。为了使仪器仪表的用户，生产、科研、管理部门能够从大量的产品信息中查询到本部门所需产品的技术经济信息，用以满足用户产品选型、科研开发、管理等方面的需要，便于宏观管理和微观信息导向，我们利用MINISIS建立了仪器仪表产品数据库。

为了满足以上各部门的需要，在调查研究的基础上进行了系统分析，我们确定了产品及生产企业的有关信息；包括产品的用途，主要技术性能和技术指标、工作原理、结构特点、关键技术、新技术、成套性、可靠性、工作条件要求、系列化、标准等级、开发形式、水平年代以及销售量、价格、成本等40个数据项。

二、数据组织和处理

仪器仪表产品分为13大类，146小类，776个系列，17000多个品种，10万个规格。其分类特点是：各大类划分小类的标准各不相同，有的根据被测变量系统的功能或成套性划分，如工业自动化仪表；有的根据产品的用途和使用方法、应用范围等划分，如电工仪器仪表，有的又根据产品的结构原理及应用领域划分，如光学仪器。同样，小类划分为系列、系列划分为品种的依据也各不相同。

仪器仪表涉及多种学科和技术，如微电子技术、信息技术、微波技术、

激光技术、传感技术、遥测遥控技术，以及电磁学、光学、声学、热学、仿生学、化学、海洋学、地球物理学、控制论、系统工程理论等。所以产品的功能及技术参数中涉及很多SI单位符号，这个数据段中的文字、数字及符号混杂在一起。

面对这种非文献型的、并要满足不同用户需要的复杂数据，如何组织和处理数据才能即满足宏观统计又满足微观查询的需要呢？

一物一码是最好的办法，但由于在国标、部标中没有这种编码，而目前国内使用的全国工农业产品(商品、物资)分类与代码(GB7635-87)中的产品编码只标到系列一级，截至到1985年统计系列下的品种数最多达6000种，规格数最多达19000多个。显然，只采用产品编码是不行的。在系统分析的基础上，根据仪器仪表产品的特点，我们利用MINISIS的功能，采取了分类法和主题法相结合的形式，系列分类及品种划分采用代码形式，主题法利用MINISIS的抽词功能，将有关字段中的关键词及技术数据抽词。用分类法进行宏观查询，用主题法进行微观查询。

主数据库中，有三个字段要进行抽词倒排，而且内容各不相同。包括产品的用途、主要技术性能和技术指标、关键技术和新技术。产品的用途、关键技术和新技术两字段的抽词相当于自由词标引，而主要技术性能和技术指标的抽词是一件复杂的工作。因为仪器仪表的技术参数要牵扯到各个学科的物理量，内容繁杂，抽词方法不同，直接影响倒排文件的大小、检索速度。通过对大量数据进行分析，我们找到了一个可行的抽词法，包括标识符、参数的选择、规范化等，在实际应用中效果很好，并在进一步完善。在其余的字段中，有三个字段采用了国家标准编码及规范化写法，一个字段采用了部信息司的编码，一个字段采用国家科委的规定写法，两个字段自行定义规范化写法。

三、KSAM库的作用

在主数据库中，有产品生产企业名称、地址、电话、电传、电报挂号等有关信息。同一个企业，有多种产品品种或规格，为了避免重复输入这些相同信息，降低由于重复输入而引起的错误，所以建立了企业基本情况数据库(KSAM库)，以企业名称代码为主关键词，将这个库作为主库中的企业名称代码的合法验证库。在主数据库中，只将生产产品的企业名称代码输入，其他有关企业的信息不输入，从KSAM库中就可获得。

建立KSAM库，对于国外产品信息的输入显得更为重要。我们搜集了美、英、日、法、苏联、西德、东德等十几个国家的产品，其生产企业涉及到多种文字，输入的错误率更高，建立KSAM库后，不但解决了这个问题，而且解决了同一企业有多种称呼的问题。

由于建立了KSAM库，不同产品的同一企业的有关信息，不用重复输入，提高了工作效率。

四、结束语

仪器仪表产品数据库采用汉字抽词倒排检索的方法,弥补了国内为计算机检索而制订的现有物资编码不能进行微观查询的不足。为了保证产品数据库的源、渠、库、流，我们将组织相应的信息网络，不定期更新数据，随时对外查询。

仪器仪表产品数据库作为仪器仪表技术经济信息系统的基本数据库之一，将会得到不断的发展。

Application of MINISIS to Family Planning Material Retrieval

Liu Zhao-hui
Yoa Xiang-yang
Sichuan Population Information Centre

MINISIS 在计划生育资料检索系统（FPMRS）中的应用

四川人口情报中心计算机室　刘兆辉　姚向阳

【提要】　本文主要介绍了利用 MINISIS 数据库管理系统建立计划生育综合资料检索系统的设计思想、功能实现和其中部分 MINISIS 数据库，并对该系统的维护和安全做一般性介绍．

[abstract]，

In the paper, the writer makes clear the design for having made good use of MINISIS database management system to creat FPMRS, functions of the system and some databases. At the end of this paper, the maintance and safety of the database have been described.

一、　概述

四川人口情报中心成立于一九八四年，是直属于省计划生育委员会的一个区域性情报科研机构．她担负着在西南四省、区范围内提供人口情报的收集、交换、检索服务工作的任务，为计划生育部门和人口研究机构提供资料和数据，为更有效地发挥已有人口情报资料的作用，及时、全面地满足各界对计划生育方面的资料需求，我们在 IDRC 提供的 MINISIS 数据库管理系统的基础上研制并建立了四川省计划生育综合资料检索系统(FPMRS)，它全面、综合地反映了四川省各地、市、州的及县(市、区)计划生

育系统的工作情况和各种指标的数据构成。为该类资料的查询提供了计算机检索服务。

二、 FPMRS 系统设计和功能

本系统的数据内容由六大部分组成，每部分均为相对独立的 RD（关系定义）逻辑结构。分别从各自不同的侧面反映了计划生育工作的基本情况。由于这六个部分的内容相关性不大，故在系统设计时将其分立为六个独立的 MINISIS 数据库。它们是：1. 四川省行政、自然区划；2. 出生婴孩按孩次分布及妇女初婚情况；3. 节育补救手术情况；4. 育龄妇女及已婚妇女现有子女分布、构成；5. 已婚育龄妇女落实节育措施情况；6. 公安系统人口变动分类对比情况。其文件类型均定义为 MASTER 类型。对于其中的主要指标数据及用户普遍感兴趣的内容进行投影操作，形成各自的 PS（投影子集）。为了在对某一部分资料实施检索的同时能参照、对比其他部分的资料，我们又将由六个 RD 产生出的 PS 连接起来而形成一个新的逻辑关系，组成了 TOTAL 数据库，即一个 DS（数据子模）。（见系统逻辑结构图）。

MINISIS 有许多特点尤其是具备建立事实数据库的特点；它能处理可变长记录和字段、允许有九个子字段，每个记录可包含 250 个字段，而每个字段可长达 4000 字节，则每一个记录总长可达 17000 多个字节。这些特点能满足我们建立记录级别较高的事实数据库的需求。MINISIS 还能建立各种结构先进的倒排文件、响应速度快、检索途径多、具有数值计算和集函数的功能，这些特点又有利于应用人员对资料的检索、对比、加工，而且 MINISIS 本身的功能齐全、数据定义方式灵活、屏幕对话简便，还能将某个数据库分解为若干个子集供若干用户分享，或将若干数据库结合成一个新的逻辑结构。我们针对数据检索的特点和要求，分析 MINISIS 的关系型数据库及联机情报检索方面的功能，确认 MINISIS 完全能够满足和支持我们系统的要求，尤其是面向一般用户更显的方便、清晰、对话友好。

在定义六个 MINISIS 数据库时，我们针对字段标识符做了如下的约定：FIELD NAME（字段名称）全部定义为汉字字符集，并自动提示；所有的 MNEMONIC（助记名）都以字段名称的汉语拼字打头字母缩写，以便于使用人员的辩别和操作。

第一个 RD 型数据库 QH（行政、自然区划数据库）中，我们分析了区划查询的内容和一般要求，确定了记录中 14 个字段的属性。它们是：区划时值（A100）、地区名（A10）、县（市、区）总计（A120）、其中（A130）、区公计总计（A140）、乡总计（A150）、镇总计（A160）、街道办事处总计（A170）、市区项（B100）、平原项（B110）、丘陵项（B120）、山区项（B130）、高原项（B140）、备注（B150），其中有 7 个带子字段的字段。由于在一个省辖的区域内市（省辖市），地区、州（民族自治州）的名称是唯一的、不重复的，因此我们定义地区名（A110）为查重项，对出现重复的地区名称，系统会提示查重信息，避免在“ENTRY”（联机输入）模块中的误操作，并定义为倒排项，使用户能通过这一使用频率

很高的字段迅速找到有关记录。还定义字段其中(A130)的子字段县总计(A131)为倒排项，满足以定量分析为主要目的快速检索需要。针对资料的必需程度和复盖率等特点，QH 库中选择了地区名、县(市、区)总计、乡总数三个字段产生一个用户要求的子关系，建立 PS(投影子集)。并在头标信息的 WRITE ACCESS TO PS? 中定义为 N，不允许检索者通过该 PS 对物理文件执行写操作。保证物理文件的正确性和唯一性。同时在回答 RENAME RD TAG 时为 N，使 PS 字段与 RD 字段的 TAG(字段标识)一致。通过定义适当的倒排项和选择访问 RD 或 PS，使资料的检索迅速，可简可繁，点面结合。

对另一个 RD 型数据库 CSYH(出生婴孩数据库)的定义也是突出方便查询的特点，并同样限制用户写操作。根据出生婴孩按孩次分布资料内容的分类，我们确定了记录中 10 个字段的属性，即时期(A100)、数据范围[1 全省、2 城市、3 镇、4 农村、5 分县](A110)、地区名(B100)、出生婴孩数(C100)、★一孩★(D100)、★二孩★(E100)、★多孩★(F100)、计划生育(G100)、女初婚(H100)、备注(I100)。为更清楚地反映孩次分布的状况，在一孩、二孩、多孩项下各带有 1. 孩次绝对数、2. 孩次%、3. 其中计划内绝对数、4. 占出生总数%、5. 其中计划外绝对数、6. 占出生总数%等子字段。与之相关因素的女初婚项(H100)，也带有 3 个子字段说明状况，即女初婚人数(H101)，其中 23 周岁以下初婚人数(H102)、晚婚率(H103)。在该数据库中，地区名(B100)是唯一的，不可重复的，故将其定义为查重项，同时建立倒排文件。由于资料的内容构成有全省总计、分地区总计、城市总计、城镇总计、农村总计、分各县总计的不同的统计口径，为了明确区分数值反映的范围，在 A110 字段中将其名称定义为可选择型的形式，数据输入人员只需在该字段的提示下打入 1～5 阿拉伯数字，便很容易地规定了记录的反映对象，也避免了检索时产生混淆。还定义了一孩率(D102)、二孩率(E102)为倒排项。用户可通过这三条倒排字段实现多途径、快速检索。在 DBCSYH 的 PS 操影操作中，选择了地区名、一孩率、二孩率、计划生育率、晚婚率、出生婴孩数六个字段构成新的逻辑子关系，也规定不改变原 RD 的 TAG，不允许通过 PS 对 RD 进行写操作，确保该库的原始数值不被非法修改。

对其余的 4 个 RD 型数据库及各自的 PS 的设计，都是基于上述原则进行的，使之都具备相同或相似的结构和功能。

本系统的设计突出和强调检索方式和检索结果处理的功能(见系统功能组织图)。检索者根据所需资料的详简程度，有选择地访问 RD、PS 或 DS 数据库。每一用户进入 QUERY 方式后，都有若干途径可供选择来进入检索点。可采用内容检索，也可采用结构检索，还可以两种方法组合使用。若我们知道某一数值的记录内部序号，则可以直接按照 ISN 进入检索。若我们重点在于查全率，则可采取自由文本检索的存在、空缺两种方式进入检索，这实际上是一种顺序检索的方式，速度较慢。若是着重于查准率和速度，就可以利用倒排文件选择关键字、关键词和关键字段进入快速检索。当然，前提是仅对建立了倒排文件的数据有效，这样考虑，既照顾了不同用户检索习惯的差异，又扩大了检索入口。在此过程中，检索者还可通过使用 QUERY 的其他命令来显示倒排字

段标识和选择隐含查询字段，利用布尔运算可以临时修改检索策略，随时扩大或缩小检索范围。由检索产生的命中记录，根据检索者的各种需求也采取多种方式的结果处理。首先，用户可在终端显示当前的检索结构，如认为对以后的检索有帮助，则可以将其储存起来以备使用，否则放弃。也可以寻找上次运行 QUERY 时存储的一个命中记录，或显示参照，或与其他命中记录组合形成一个新的检索结果并以不同方式将其排序，经过排库处理后的记录，可送到终端显示，也可按照隐含格式或特定的指定格式输至打印机打印出来。还可以将其在文件中保存起来供以后参考和使用。如果还要做进一步处理则将检索结果直接输出到其他的处理程序（PRINT、INDEX、COMPUTE 等）。通过这些功能的选用，基本上可以满足不同的服务对象和服务要求。

该系统的最后一部分是系统维护，它是整个数据库正常安全运行并不断改进其性能的保障。MINISIS 提供的一组服务程序，帮助我们十分便利的进行空间调整，逻辑重构造、数据库转贮和系统恢复等工作。我们还采取了一系列的保密措施，以保证数据库的安全，防止非法用户的访问。在系统中的用户分为检索者，输入员、管理员三类，并设置了各自的口令和能力，贮存在用户保密户头文件中。我们规定了检索者的能力仅限于使用 QUERY 和 PRINT 模块，输入者可增加运行 UPDATA 功能模功，而数据库管理员则可运行 MINISIS 所有的功能模块。并严格、慎重地提供第二类用户的用户名，以确保数据库的安全。

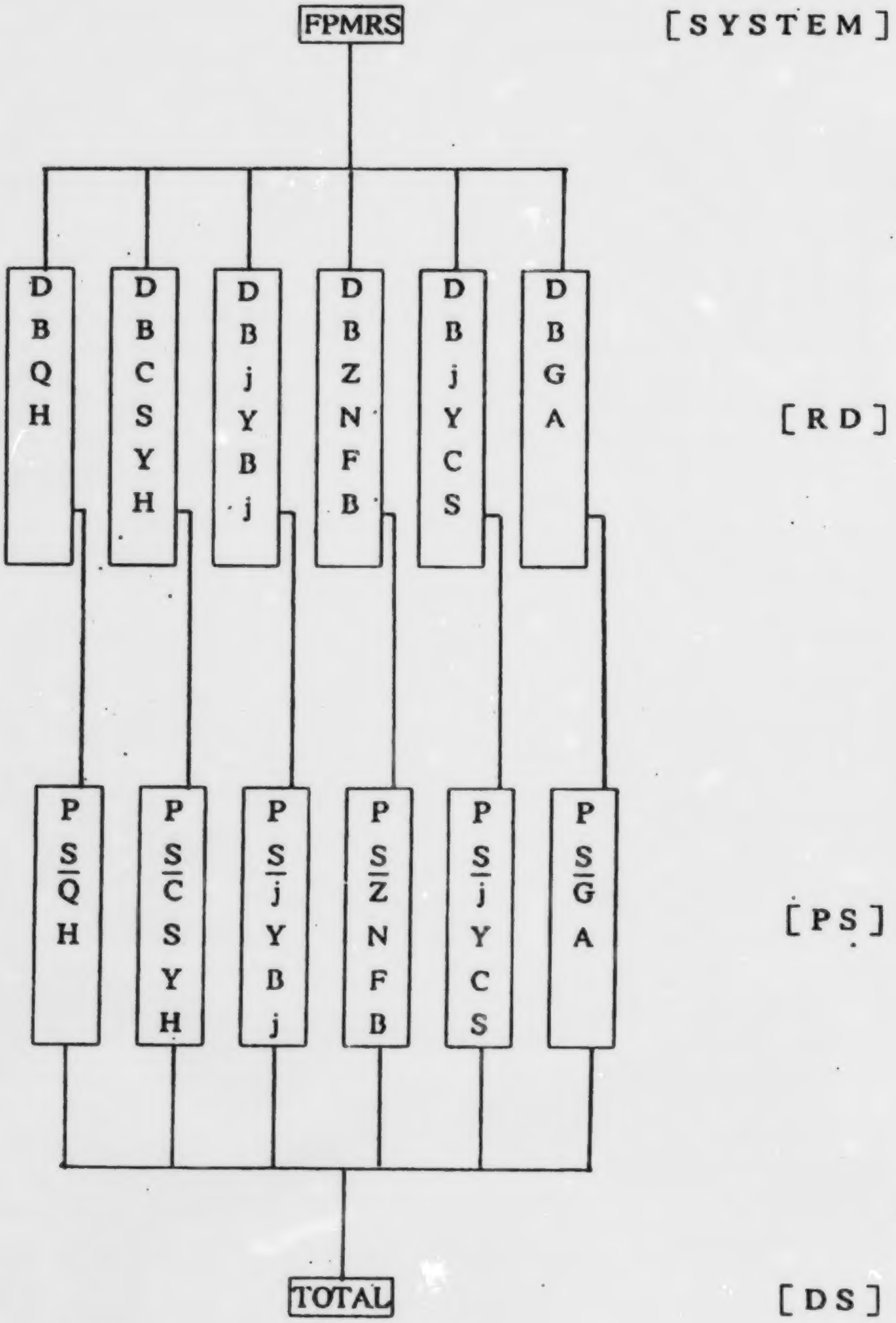
三、 小结

本系统是面向检索者并努力为满足其检索要求而建立的检索系统，因此，在设计中着重突出了检索入口宽、途径多、速度快和结果处理多样化的特点。由于有 MINISIS 强大的数据处理和情报检索的功能支持，使得我们的系统研制得以初步成功。在试运行中已显示出它的许多优越性，尤其是一些手工方法难以做到的比如条件查询、运算处理等工作，在本系统中已能较容易、准确的实现。

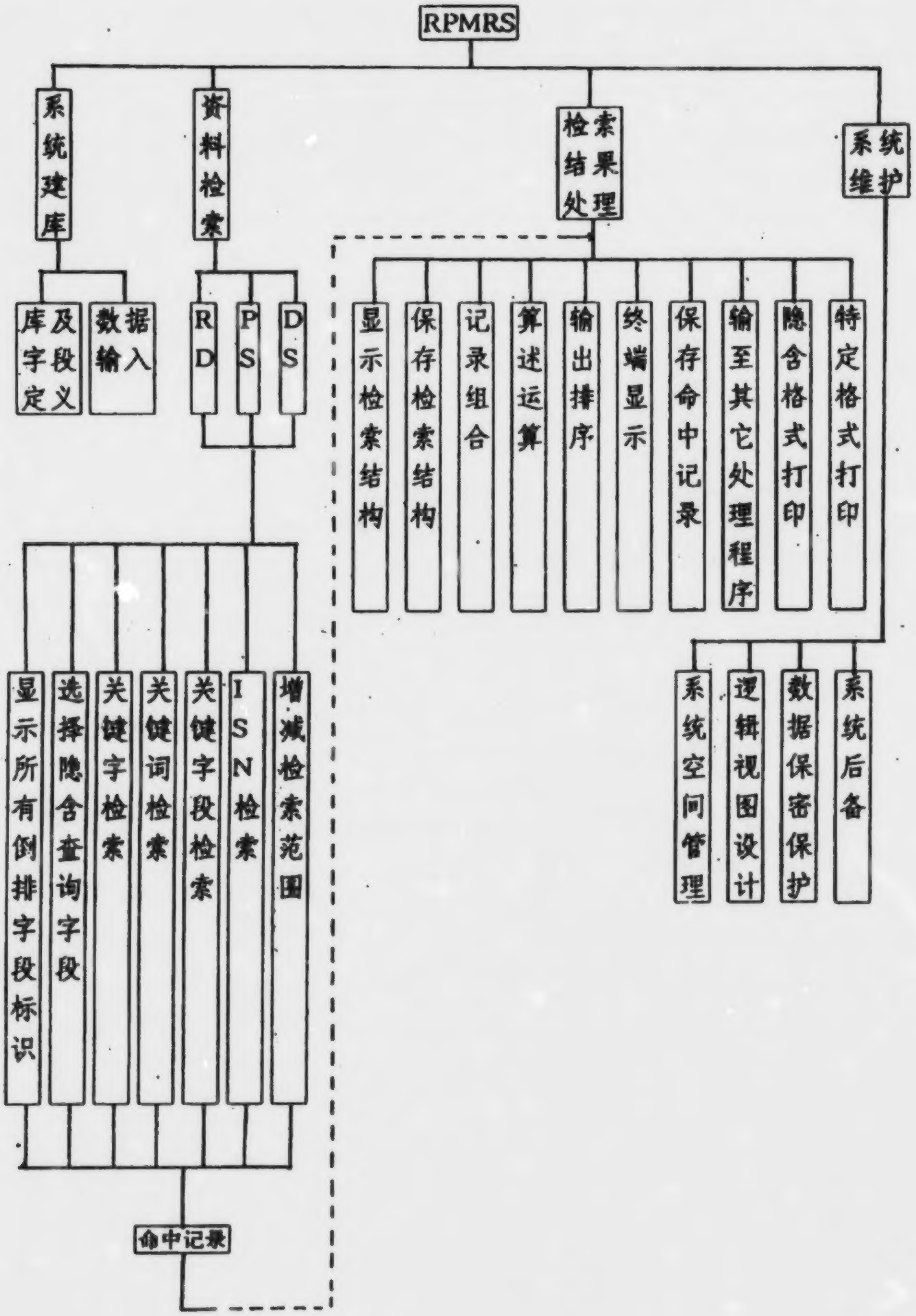
本系统目前尚不尽完善，待进一步改进后投入正式运行。

1989 年 10 月

系统逻辑结构图


FPMRS
[SYSTEM]
DBQH
DBCSYH
DBjYBj
DBZNFB
DBjYCS
DBGA
[RD]
PSQH
PSCSYH
PSjYBj
PSZNFB
PSjYCS
PSGA
[PS]
TOTAL
[DS]

系统功能组织图


RPMRS
系统建库
资料检索
检索结果处理
系统维护
库及字段定义
数据输入
RD
PS
DS
显示检索结构
保存检索结构
记录组合
算述运算
输出排序
终端显示
保存命中记录
输至其它处理程序
隐含格式打印
特定格式打印
显示所有倒排字段标识
选择隐含查询字段
关键字检索
关键词检索
关键字段检索
ISN检索
增减检索范围
命中记录
系统空间管理
逻辑视图设计
数据保密保护
系统后备

The Establishment of Computerized Agricultural Databases in Chinese Language

CHUNPEI HE
Scientech Documentation and Information Center
Chinese Academy of Agricultural Sciences

I. Background:

The Scientech Documentation and Information Center (SD lC) of the Chinese Academy of Agricultural Sciences has introduced H P3000 37 minicomputer system in 1987 under the kind donation by lDRC. Since then, SD lC has successfully established CABl and AGRIS databases by using MlNlSlS developed by lDRC and the service of both selective dissemination of information and retrospective searching has ever expanded. However, the establishment of our own databases in Chinese language is still a new issue for our colleagues who have used MINISIS for just one year or so. The leader of SDlC allocated the assignment of establishing Chinese Agricultural Databases with the objective of serving local users by our own library collection. This is of vital importance because many users complained about not having enough literature support from the local library after searching the foreign databases.

II. Databases under establishment:

SD lC is now establishing three databases in Chinese language: (I) Chinese Agricultural Abstract; (2) Chinese Agricultural Literature Citations; and (3) Agricultural Research Achievements database. The first one is similar to CAB I, and the second to AGRlS. SDlC is planning to establish another database: Current Research Project Database which is similar to CARlS. But due to the financial issue, the work might be postponed until I 990.

III. The Methodology of Establishing These Databases:

We used MlNlSlS for establishing these databases in September 1987. After 6 months of experiment, we found input the data into the computer via the H P3000 terminals is a kind of waste in terms of both electricity consumption and time. Especially when we use HP3000 37 with HP7935 disk drive, the temperature restriction hinders our speed in the establishment of these databases.

In order to remedy this issue, we introduced Micro CDS ISIS in inputting the raw data on microcomputers or Asian Vectra. In so doing, the inputting work could be done locally on work stations independent of the status of HP3000 37. What we do is to convert the input records into ISO 2709 format from microcomputers and transmit them into HP3000 37 through Advlink software from Asian Vectra. The production of inverted files for searching is to be done on M IN 1515. The successful experiment has enhanced our confidence in the establishment of these Chinese databases.

IV. Agricultural Research Achievement Database:

For the Agricultural Research Achievement Database, we have used more than 30 fields. This database is very similar to the CAR IS of the Food and Agriculture Organization

(FAD) with the addition of many pieces of useful information including Chinese Classification Scheme which is very important to the local users in query. There are several repeated fields and some of them have subfields within them. The users can search the database by participants of the research, subject category, descriptors, keywords, keywords in title, date of prize awarded, etc. Due to the fact that the Chinese language thesaurus has not yet been completed, we set up a field to let the cataloging people enter free terms. In order to let the users search title words, we also set up a field called keywords in title. In addition to the essential elements in the database which include a summary of 500 to 800 Chinese characters, we also include economic benefits, social benefits, ecological benefits and scientific value of the research achievements. This database now contains 2,000 records of research projects which have won prizes from the Ministry of Agriculture and from the State Science and Technology Commission from 1978 to 1988. After one year of experiment, the users are very satisfied with the rearch result on this database.

V. Chinese Agricultural Abstract:

The Chinese Agricultural Abstract database is very similar to CAB I. We set up more than 25 fields including many pieces of essential information. The structure is somewhat similar to AGRIS database established by FAO. However the content is somewhat similar to CAB I. This database includes a field entitled abstract which contains a text of abstract of about 300- 400 Chinese characters. In addition to this, we include descriptors, subject index elements, and free terms. The keyword in title has further enhanced the search benefits for the users. The wonderful thing we do is to have two kinds of output. In other words, from one inputting we can establish the computerized database in H P3000 minicomputer and also have the needed fields to be published in abstract journals which will be circulated throughout the country via post offices. This database now contains more than 2,000 records in horticulture. The following year, 5 other disciplines will be included in this database. They are field crops, soil and fertilizer, animal husbandry, veterinary medicine and plant protection. The number of records in this database will be increased by 10,000 annually from 1990 on.

VI.Chinese Agricultural Literature Citations Database:

The Chinese Agricultural Literature Citations database is a comprehensive one. This database is very similar to AG RIS. It contains all sorts of literature within our library's collection including information from monographs, journals, reference books, serials, proceedings, and patents. Up to now, this database includes more than 20,000 records covering literature of the recent two years. We set up about 15 fields in this database. We have introduced the subject category of AG RIS into this database with the ojective of enhancing the query results from computer searching though we already have the Chinese Classification Scheme. What we intend to do is also to kill two birds with one stone, namely through inputting the data into the computers we can obtain two kinds of output: the establishment of the database in minicomputer for future query and the publication of journals. The publication of these journals is done by converting the data from the computer into 150 2709 format and further composing them on another microcomputer system. The laserjet series I I is hooked to one of the Asian Vectra for computerized composition. The printing is of very good quality.

VII. Problems and Discussion:

It is a successful technical route to use Micro CDS 1515 to input data into microcomputers and transmit the data via 150 2709 format into M IN 1515 in the establishment of Chinese language databases. However, during our experiment. we have found that when we use handler in the G version of M INISIS, there is a problem in using VPLUS. If the handler is not used, there is no problem in using VPLUS. If the handler is used, some of the data in the table will be lost when we use VPLUS tables to modify some of the fields of a record.

On the other side of the scale, the sort table in M IN 1515 could be used if we use handler. This is a very good feature of M IN 1515 for our institute, because we could produce an annual index of either author or subject in the last issue of each journal. This can remedy the shortcoming of "GB 2312 80" which classifies Chinese characters into two classes, with Class One sort according to pinyin and Class Two sort in accordance with the strokes.

Last of all. we hope that the sort table could be further improved in the following aspects:

(1) The sequence of the Chinese characters with the same pinyin could be rearranged according to Xinhua Dictionary.
(2) Solve the issue that some of the Chinese characters could not be entered into the computer.
(3) A further improvement of the printing format of the sort table.

MARC Applications in Chinese Bibliographic Databases

Ms. Wang Ming
China Population Information Centre

MARC 在建立中国人口情报书目数据库系统中的应用

汪　明

引　言

本文阐述了在建立中国人口情报书目数据库系统的过程中，利用MINISIS数据库管理软件开发MARC应用系统。为了保证中国人口书目数据的标准化，便于交换，在安装MINISIS G版本后，我开始了对该系统的研制工作；用MINISIS软件处理MARC书目数据，特别是对中文书目数据的处理。在此，向大家作一回报，希望得到各位，尤其是有关专家的指点与帮助。

首先，我十分感谢IDRC及有关的MINISIS用户在如何使MINISIS系统适应MARC格式数据要求方向所作的大量研制工作，能够提供MINISIS–MARC接口，给其它MINISIS用户带来了很大的方便。在考虑采用MARC数据交换格式的同时，我对MINISIS系统为接受MARC格式数据在各方面进行的补充作了一定的研究及尝试，我认为，经过接口程序的设计，使MINISIS数据库管理系统能够更理想地处理书目数据，并且可以它是一个非常实用的书目数据库系统的管理工具。

另外，我可以告诉大家，在目前世界上众多的MARC形式当中，我们中国也有了自己的MARC即CMARC，为我们的中文机读目录数据提供了标准的依据。它的基本标准是《国际机读目录格式》“UNIMARC”，其中作了一些必要的补充和修正，以适应中文数据处理的需要。现在，我国国家书目机构正准备以该格式向国内提供中文机读目录数据。因此，用MINISIS进行中文机读目录数据的处理及利用有了充分必要的保证。

在坐的各位代表，你们其中有一些专家在利用MINISIS处理MARC数据方面的应用已经有了很多的经验，我很希望能够得到你们的帮助，尤其是在LCMARC应用方面。然而，作为中国用户，特别是我们如何利用MINISIS对CMARC数据进行处理，还是一个新的起步。因此，我就着重地向大家汇报一些我们在CMARC数据处理方面所做的实际工作，但首先，我简单地对CMARC和MARC接口作一概述。

CMARC 简介

目前，我所介绍的只是CMARC中有关图书机读目录格式。依据UNIMARC规定，凡是适用于中国出版物特征的UNIMARC的有关规定均予遵循。对于中国出版物特有而UNIMARC中未做规定的，CMARC加以规定。

CMARC参照的中国标准有：

《中华人民共和国行政区划代码》
《信息交换用汉字编码字符集第二辅助集》
《信息交换用汉字编码字符集第四辅助集》

《信息交换用汉字编码字符集第一辅助集》
《信息交换用汉字编码字符集第三辅助集》
《信息交换用汉字编码字符集第五辅助集》
《中文书刊名称汉语拼音拼写法》

由于时间关系，对 CMARC 增加的内容就不多做说明了，具体使用时基本的形式是一致的。

MARC 接口概述

由于 MINISIS 是一个通用的数据库管理系统，适用于多种文字数据的管理，尤其是在文献、书目资料的管理及应用方面具有它的特长，因此，可以利用它来作为书目数据处理的软件工具。但为了适应 MARC 数据格式的需要，在用 MINISIS 数据库实现对 MARC 数据结构的定义时，需要对原来的 MINISIS 数据库定义作一些必要的补充及修订，满足 MARC 记录中数据单元与 MINISIS 记录中字段的一一对应关系；并且需要编写出相应的处理程序，从而完成对补充或修订的非 MINISIS 格式数据的处理。为此，我们需要利用 MINISIS 提供的"接口"手段，将这些所作的接口软件连接到 MINISIS 软件上，实现对 MARC 格式数据的处理。

在根据 MARC 定义的数据结构格式建立 MINISIS 数据库的同时，接口程序必须将 MARC 结构的数据单元映射为 MINISIS 数据库定义中的字段，同样，经过接口程序的处理，可将 MINISIS 数据库字段映射为 MARC 结构的数据单元，从而能使 MINISIS 对 MARC 数据进行处理。为了协助这种映射处理，需要建立一个文件，以存储映射处理需要的信息，该文件可作为一个参数文件与 MINISIS 数据库相连，任何一个或所有接口过程因此可存储这个文件，以获得我们所定义的 MARC 结构信息。这个文件即是对 MINISIS 数据定义的补充，相当于 MINISIS 字段与 MARC 数据结构单元的对照表。当 MINISIS 内部过程需要存取这个 MARC 数据结构的物理文件时，它调用 2 个 MINISIS 全局出口程序中的一个出口程序，这个出口程序本身并不进行任何处理。每个出口程序简单地调用为处理 MARC 记录而编写的特殊过程，并通过这个过程执行所需要的接口处理。

总体上说，建立 MARC 接口需要做很多工作，定义 MINISIS 数据库；建立参数文件；修改全局出口程序；编译接口并存入 MINISIS 的 SL 文件中；编写处理 MARC 记录的过程，并对其进行编译，放入 MINISIS SL 文件；建立 MARC 数据格式所用的物理文件。

一旦接口准备就绪，对以 MINISIS 数据库定义的 MARC 记录的处理与 MINISIS RD 的使用方法一样，记录可以输入、修改、打印、删除。倒排文件将按数据库定义中每个字段的倒排参数进行更新，记录可以用倒排文件进行检索，亦可将 MARC 格式数据库当成 RD 那样，建立投影子集或作为 DS 数据库中的一个成分。除此之外，使 MINISIS 系统具有了新的特点：

1)能够对 ISO2709 的头标区数据进行访问;
2)数据库记录的子字段可以多于 9 个;
3)母字段中个别子字段的重复;
4)对字段中某固定位置上数据元的处理;
5)排除数据中带有子字段标识符时给排序、检索或打印时带来的麻烦。

这样就为保证 MARC 记录的处理提供方便的手段。

下面，将详细介绍在建立中国人口情报书目数据库时，对 CMARC 的应用。

CMARC 系统的应用

中国人口情报书目数据库系统分中、西文两个独立的子系统。中文书目数据采用 CMARC 格式，从而保证了标准书目数据的来源与交换。在此，以 CMARC 为例，介绍一下我对 MARC 接口系统的应用。

一、建立 *CMARC* 应用环境

为了能够使 MINISIS 处理中文机读目录 CMARC 记录，必须在原有的系统下配置汉字接口系统和 MARC 接口系统，满足中文机读目录格式数据的要求。

设置汉字接口系统

汉字的处理，需要使原来的 MINISIS 系统扩充标准字符集，而采用不相似字符集，并在原始的 MINISIS SL 和用户帐号 PUB 组的 SL 上增加内容，将代码转换表、汉字字符集属性表 CSAT、汉字字符集定义过程 CSD、常用汉字字型和汉字排序表及所有管理程存储序到 MINISIS SL 或用户帐号 PUB 组的 Sl 中。

目前我选择了 GB2312 作为内部编码，使用了 MINISIS 系统中 MINLIB 帐号的 LANGCHIN 组中提供的原始代码转换表、所有的管理程序、汉字字符集属性表 CSAT、汉字字符集定义过程 CSD，通过输入输出设备的管理程序，定义了自己的输入方法。我们的输入终端有中文终端 HP2693A、亚洲工作站 Asian Vectra，能处理 6763 个汉字。输出采用伪脱机方式，通过 LOADGB 作业还将常用汉字字型和汉字排序表装入了 SL 中。

总之，设置汉字接口系统，实质上需要保证在所运行 MINISIS 系统的帐号下，PUB 组的 SL 中必须包含以下各段:

段名：CHIN'LANG'TOOL 见附表 1)

段名：GBTELECODEI

段名：GET'TELECODEI
段名：CS'PATTERN

设置 MARC 接口系统

以上已经在 MARC 接口概述中大致谈到了建立该接口的工作内容。现就我们实际工作的过程及所发现的一些问题，向大家作一次汇报。具体分五个过程：

一、建立 MARC 接口所用参数文件中的参数

作为 MARC 接口自己的信息集需将"MARCMES0"加载到 MINISIS 系统的信息文件上，供 MARC 接口程序调用。MARC 接口的参数文件内容的起始地址应在 MESS00 文件中第 2294 行上。

建立方法：

```
EDITOR
/T MESS00.PUB.MINISIS
/J MARCMES0.MESSAGE,MINLB
3149...



/ R2294
  2294     3149   \
/K MESSOO.PUB.MINISIS
/EXIT
```

MARCMES0 信息集内容参见附录 1。

二、建立 MARC 定义参数表 MARCTBLE

在以上的概述中已经谈到了为了补充对 MARC 记录的定义，类似于 MINISIS 中数据定义文件，可以利用 EDITOR 兼容格式的文件形式将 MARC 记录的映象如子字段、固定字段的处理定义进行补充。

在建立 MINISIS 数据库时，当头标区出现"NAME OF PARAMETER FILE"提示时，提供一个合法的 MPE 文件名，并存储在 MINISIS 数据库定义中，可由 MARC 接口过程检索来处理 MARC 记录。它是一个未编号的 EDITOR 文件，其中每行存储着由 MARC 接口所要求处理的固定字段或子字段的有关信息 255 个字符。由该文件记录中的不同位置上的参数表示不同的

处理意见.

MARCTBLE 的建立有两种方法. 直接对 EDITOR 文件进行编辑:

```
EDITOR
/Set length=255,right=255
/a:





/K<名>,UNN
/E
```

利用 MINISIS RD 生成: 用数据库中的字段内容代表 MARC 定义参数表中各列内容、由打印程序将此库转化为 MARC 定义参数表.

数据库: MPE 合法名

MINISIS Tag
Subfield /Fixed field identifier
field name
field offset
field Subfield prompting?
Sort key generation?(y/n)
Subfield printing?(y/n))
Inverted key gengration?(y/n)
Word processing of Subfield(y/n)
primary /2nd key extraction(1.12)?
prompt for field if MONOGRAPH?
prompt for field if ANALYTIC?
prompt for field if COLLECTIVE?
prompt for field if SERIAL?
Right justify fixed field?(y/n)
Fixed field fill charater

一旦按各字段输入相应的参数值后, 就可利用 IDRC 提供的打印格式 PMARCT00 将其输出为一个 ASCII 格式文件, 作为 MARC 接口所需的定义参数表, 并能经过 EDITOR 修改该表中内容.

三、安装 MARC 接口

MARC 接口包括全局出口程序和局部接口过程。全局出口程序永久地与 MINISIS 软件连在一起，常驻 MINISIS SL 中。由 MINISIS 内部过程调用；局部接口只从事特殊工作，并针对一个指定的数据库。存在于非 MINISIS 帐号的 PUB 组或用户的联机组中。

MARC 接口由九个出口程序组成，分全局及局部两部分。全局接口作用为：产生倒排或检索所使用的关键字、作为显示信息所要求的子程序。IDRC 提供的 MARC 全局接口有：

USER'GENKEY：它是一个哑过程，用于调用 GENKEY'FUNC1 子程序。
GENKEY'FUNC1：字处理倒排、MARC 子字段数据倒排、用于功能代码为 1 的抽取方式。
USER'DISPLAY：它是一个哑过程，作为调用从事 MARC 子字段打印的子程序。
MARCPRNT'FUNCI：作为按某格式显示 MARC 子字段的子程序。
LOAD'MARC'TABLE：该子程序的功能是从 MARC 表中获得信息并以可读形式存放于内存中，可由前几个过程调用。
MARC'FLD'ATTR：从已装人内存中来自 MARC 表中的信息获得一个已给定的固定字段元素或子字段的性质。

局部接口有：

MARCINP IPFIXED：对 MARC 固定字段数据输入、更新的处理。
IPSUBFLD：用于控制子字段数据的输入。
MARCINDX：在 INDEX 中生成排序键。
LOAD'MARC'TABLE：从 MARC 表中获得有关信息并以可读形式放在内存中，此过程可由前三个子程序调用。
MARC'FLD'ATTR：从正装人内存中来自 MARC 表中的信息得知一给定 MARC 固定字段或子字段的属性。

使用 MARC 接口必须将 MINISIS SL 中两个哑程序用该接口代之。即重新修改 MINISIS SL，增加一个段 MARC'ROUTINE，将已编译好的接口 MARCEXIT.OBJECT.MINLIB 安装在此段中。

过程如下：

```
编 译 接 口: SPT MARCEXIT, MARCXUSL
取代哑程序:  SEGMENTER
并增加 MARC:   SL SL.PUB.MINISIS
接 口 程 序: PURGESL ENTRY, USER'DISPLAY
             PURGESL ENTRY,USER'GENKEY
             USE MARCXUSL
             ADDSL MARC'ROOTINE,PAMP
             EXIT
```

这样一来，在 MINISIS 帐中的任何用户，都可以使用本帐中 PUB 组的 SL，当调用 USER'DISPLAY 或 USER'GENKEY 时，只执行以上的局部出口程序。

如果使用非 MINISIS 帐号，用户就不能访问局部接口。为了保证能在其它联机帐或组中访问 MARC 局部接口，简单的方法就是将已修改过的 MINISIS SL 复制到所使用的联机帐下的 PUB 组。过程如下:

```
:FCOPY
FROM=SL.PUB.MINISIS;TO=SL.<组>.<帐>;NEW.
EXIT
```

目前我所使用的是 IDRC 提供的原始 MARC 接口 无须对其重新编译，并将修改过的 MINISIS SL 全部拷贝至非 MINISIS 的联机帐中。

但如为了节省空间，可将被拷贝的 MINISIS SL 中无需的某段删除，或选择性地编译 MARCEXIT 文件中一部分子程序于 USL 中，并将其拷贝到非 MINISIS 联机帐下的 PUB 组。

CMARC 数据库的建立

建立一个能够处理 CMARC 记录的书目数据库，同常规的数据库建立的形式有所不同，主要在于:

<投标区>

除了将库定义为 MINISIS 类型外，必须给出“NAME OF PARAMETER FILE”的参数，指定 MARC 定义参数表，从而作为数据库的补充定义。

<字段部分>

数据内容将包含子字段标识符即“＄i”[i= a...z A...Z 1...9]、指示符及 MARC 记录的数据内容。当某字段输入受控时，需调用相应的出口控制程序。

固定长字段——由 MINISIS 不可重复的基本字段表示。字段长度包含指示符、子字段标识符及数据内容的长度，并由出口程序“IPFIXED”控制输入即:

PROMPTING CONTROLLED BY EXIT >Y
NAME OF EXIT >IPFIXED
例: CMARC 的 100 字段由 B000 表示
B000: ＄a19890927d1985   k   y0chia0120G2   e

子字段化字段—在 MINISIS 数据库中有两种方式表示:

1)用 MINISIS 本身的子字段形式表示 CMARC 记录的子字段，并由出口程序“IPSUBFLD”控制数据输入。

程序"IPSUBFLD"控制数据输入.

例: CMARC 的 676 字段由 G760 表示
G761: $ a330

2)用 MINISIS 可重复的基本字段表示 CMARC 记录的子字段化字段，数据输入时不受接口程序控制.

例: CMARC 的 102 字段由 B020 表示
B020: $ aCN $ b110000

在建立中国人口情报书目数据库系统时，我们对 MINISIS 数据库中记录字段标识符同 CMARC 记录的字段标识间的转换作了相应的规定:

用 26 个英文字母（A...Z）代替 CMARC 记录中字段标识的第一位；用 10 乘以 CMARC 记录的字段标识后两位作为 MINISIS 库记录中字段标识的后两位.

如 200 字段用 C000 表示

另外，对于 ISO 2709 的头标区数据元素，给出了特定的标识符:

LO50 记录状态	LO60 记录类型
LO70 书目级别	LO80 层次等级
L170 编目等级	L180 著录格式
L190 国内发行方式	

目前，本数据库记录共计字段 48 个（不包含子字段），记录为 1702 条.

CMARC 数据处理

在 MINISIS 系统中，对 CMARC 数据进行增、删、改、显示打印、排序、倒排和转换输出，都要参照 CMARC 数据定义参照表，针对固定字段或子字段的不同形式调用接口程序作为相应的处理.

1）数据输入:

借助 CMARC 数据定义参数表自动输入指示符、子字段标识符. 按位输入固定字段数据元素或进入隐含的数据元素；对于不受控制的子字段化字段以字符串形式输入.

2）数据打印:

固定字段数据的打印--无需接口程序的处理只要在定义打印格式文件中给出以下参数:

DISPLAYCEMENT INTO THE FIELD
MAXIMUM NUMBER OF CHARATERS

在 CMARC 数据定义参数表中给出了固定字段元素的相对位置及需要打印显示的参数值.

子字段数据的打印—当不做任何特殊处理时，用 PRINT 直接输出所输入

摸块定义的子字段参数值:

FUNCTION CODE OF USER-DEFINED EXIT

做相应的处理，需要显示的子字段必须由 CMARC 字段定义参数表中的第 38 列给出"Y"值.

3) 数据排序:

当以固定字段元素作为记录的排序键时，只要利用 INDEX 摸块中的参数"DISP="给出某固定元素位置的起始值及长度，并依据 CMARC 字段定义参数表即能进行数据的排序工作.

当以 CMARC 子字段值作为记录的排序键时，可有 6 个接口程序提供使用.

i 以整个母字段或子字段集为键: 依据 CMARC 定义参数表第 36 列的取值，调用局部接口程序"WHOLE FIELD"，自动剔除指示符及子字段标识符.

```
例: KEY=TITLE,LENGTH=200,KEEP=NO,EXIT=WHOLE'FIELD
        FIELD=C000
```

ii 以母字段中的某些子字段分别作为各键: 依据 CMARC 定义参数表第 36 列的取值，调用局部接口程序"SUBFIELD".

```
例: KEY=TITLE,LENGTH=200,KEEP=NO,EXIT=SUBFIELD
        FIELD=COOO
```

iii 以被选子字段作为排序键. 同一母字段的某子字段作为主键，其它子字段作为次键. 作主键的子字段已在 CMARC 数据定义参数表的第 44 列上给出取值为"1"，并调用接口程序"PRIMARY'WHOLE"; 作次键的子字段定义的参数值为"2"，调用接口程序"SECOND'WHOLE".

```
例: KEY=PRIM,LENGTH=50,KEEP=NO,EXIT=PRIMARY'WHOLE

    FIELD=G010
    END
    KEY=SECN,LENGTH=50,KEEP=NO,EXIT=SECOND'WHOLE

    FIELD=G010
    END
    * * *
```

如果指定重复子字段的第一个值作为主键或次键，可调用接口程序"PRIMARY'SUBFLD"或"SECOND'SUBFLD".

4）数据倒排:

借助 CMARC 数据定义参数表中第 40 列上有关抽键类型的取值规定，对子字段数据可进行联机、脱机、释放倒排处理，在进行数据库定义时，还需给出:

FUNCTION CODE OF USER EXTRACTION 的值“1”.

目前，我们均采用子字段联机倒排，由程序自动地将指示符、子字段标识符剔除.

5）数据交换:

利用 ISOCONV 模块，可将数据库内容以 CMARC 形式进行交换. 为了使 MINISIS 数据库内容同 CMARC 记录建立一一对应关系，需要建立这种关系定义“CD”，在 CD 中只定义 CMARC 字段同 MINISIS 数据库记录的基本字段间的对应关系.

由于我们采用字符“$”作为子字段标识符的内容代替了八进制 37. 故需借助接口程序“CONVCHAR”在进行 LOAD 或 DUMP 操作时，对该字符进行转换. “CONVCHAR”局部接口程序需要装在联机帐中的 PUB 组. 过程如下:

```
:SEGMENTER
-----SL SL.PUB.<联机帐号>
-----USL USL.OBJECT.MINISIS
-----ADDSL CONVCHAR
-----EXIT
```

以上是我对利用 MINISIS 数据库管理软件处理中文书目数据“CMARC”的工作介绍，望得到有关专家的指正.

Acquisition, Cataloguing and Public Access Subsystems of STIC Integrated Library System

Li, Linja
Shi, Minghong
The Scientech Information Centre

This paper is about the general environment of the STIC Library, the design principles and characteristics of its integrated library system, and the database structures and functions of the acquisition, cataloguing and public access subsystems.

Description of The Central Library for Mechanical and Electrical Engineering

The formal name of STIC library, to the public, is the Central Library for Mechanical & Electrical Engineering. Formed by eight divisions of STIC related to library functions, it is the national centre for collection, storage, retrieval, and dissemination of scientific and technological information in the fields.

The library not only serves about 800 professional and technical staff in STIC, but also has responsibility for supporting information and services to all other users and units in the industry and in the related fields in China. The 600,000 monograph and serial volumes, and 6 kinds of machine-readable files form the foundation of information resources for the library to facilitate its services.

The library uses two sets of HP3000 series mincomputer with 50 terminals and about 20 micro-computers. MINISIS applications for the library's routine work have been undertaken with the HP3000 III. The multilingual integrated library system uses terminal emulators developed on IBM compatible micro-computers in support of Chinese Western languages processing.

The System Principles and Characteristics

In 1987, the library worked out its plan for developing a computer integrated library system based on MINISIS applications. According to the plan, the whole library operation is divided into six computerized subsystems: acquisition, cataloguing, reporting, public access, serial control, and circulation. Also the plan described the design principles and characteristics for realizing the system.

A brief introduction to the system design principles and characteristics is as follows:

1. Setting up two MINISIS MASTER RDs.

There are three main categories of materials in the library: monographs, documents, and serials. Since the ways for describing and processing serials are so differ from those which for monographs and documents, the library uses two separate MASTERs: one for handling serials and another for handling monographs and documents.

2. Multilingual processing

The library's collection is in many languages, such as Chinese, English, French, German, Japanese, Russian, etc. The system was designed to process materials in different languages as a whole in each of the databases, for example, monographs and documents in whatever languages will be processed in the same database. Fields were also designed to support multilingual processing, only if it would not work this way, for example, fields such as "year of publishing," "place of conference," and "type of documents" were designed to contain their data in different languages. The fields such as title, author should be considered very carefully since the ways for extracting words from titles in Chinese, and from titles in the Western are so different. Separate fields were set up for processing titles in different languages.

3. Inputting data with worksheets

There are three reasons for using worksheets input by clericals instead of online by acquisition librarians.

a. Multilingual data input is not easily handled on the keyboard, and professional input clericals should be arranged.

b. Worksheet method allows acquisition librarian or cataloguing librarian to concentrate on filling the sheets while input clericals to specialize in the skills of input different languages; and this can improve the efficiency as well as can enhance the transformation procedure from procedure the manual system into the computerized system.

c. The data input can be checked many times to improve the quality.

4. Comprehensive processing for both original documents and secondary documents

The major library holdings are documents such as academic theses, essays as well as technical reports. Most of the documents will not only be catalogued, but also be processed and reported in the formats of abstract or citation index. Therefore the system was developed in supporting of both original and secondary document processing.

5. Setting-up validation and authority files

To guarantee the quality of data input, the system was designed to set up several validation and authority files, such as status code validation, language code validation, as well as cooperation codes which will be input instead of the full names of corporate authors.

6. Using bibliographical level control

Since the system process both the original and secondary materials, the standard codes of bibliographical levels were accepted to identify the difference between each record, to provide the retrieval subsystem with another control. The bibliographical level codes also bring the convenience for editing various kinds of indexes and abstracts.

7. Supporting standardized formats

Fields in the system were designed to support outputting catalogues or indexes in the national or international standardized formats.

8. Thesaurus application

To improve searching capability of the multilingual database of the library holdings, the online Thesaurus of Chinese Mechanical Engineering is linked with the database. The thesaurus will contain Chinese main terms, and their equivalent English main terms, as well as main terms in Chinese Pinyin and in other codes which present the same terms. Since MINISIS THLOADER module supports the translation function, materials related to one subject concept, but in different languages, can be found by either a Chinese term or an English term, by its equivalent code, or Chinese Pinyin.

Database Structures of the Acquisition, Cataloguing and Public Access Subsystems

Main databases to meet the function needs of acquisition, cataloguing and public access are arranged as follows:

1. LIBRARY RD MASTER

LIBRARY is the core database of the subsystems. It contains most of the fields accessed by the routine work of library acquisition, cataloguing, public access, and even some fields related to the function of library circulation. Three types of data are processed here: data for routine management, data for describing materials and data for statistics. And this database is only used by the system manager.

2. ACQ PS

ACQ is a PS database from LIBRARY, it supports the routine work of library acquisition, and contains all the necessary fields. ACQ is used to collect materials in all the four (A.M.C.S.) bibliographical levels, except journals and periodicals; and to print out accounting sheets.

3. ORDER DS

This database covers all the ACQ fields, and vendor code associated fields as well, such as vendor name, address; and corporate author code related fields. Therefore the full names of the vendor or corporate author can be printed out with this DS. ORDER DS is only used by acquisition staff for printing out order records and claim sheets.

4. WKSHEET DS

This is another database used by acquisition staff. It contains all the fields accessed by acquisition and used by cataloguing staff to print out worksheets for cataloguing after acquisition records have been correctly modified.

5. CAT PS

CAT is a PS from LIBRARY, which contains all the fields for cataloguing and indexing all types of materials being processed by acquisitions. Cataloguing staff will fill the blanks of the worksheets left by acquisition, and then input clericals will modify and complete the records for library holdings here with this database.

6. CANAL PS

CANAL is also a PS from LIBRARY. It is used only for inputting records created from the cataloguing routine. For example, most of the records at the analytic bibliographical level, which will be created by cataloguing staff instead of acquisition staff alone, will not have relationships to acquisition records.

7. PUBLIC DS

PUBLIC is a database for public access, and is used to search the library holdings. It has all the fields CAT has, plus fields such as corporate body codes, full names of corporate bodies, and the address, etc. An initial parameter limits users' access to records. This database is specially arranged for library users and public access staff, and it supports the printing of different formats of retrieved hits.

8. CARD DS

CARD is similar to PUBLIC, but without the initial parameter. It allows to access records which are in process. This database is used by cataloguing and indexing staff to check the completed records, as well as to output various formats of catalogues and indexes; and it is also used by public access staff to read records which are not available in PUBLIC.

9. INST RD (KSAM)

INST contains detailed fields needed to describe institutes and corporate bodies in LIBRARY database, such as institute or corporate body codes, their full names, addresses, and so on. It is the authority file for corporate bodies, and will not be used alone.

1O.INSTC RD (KSAM)

INSTC contains the same fields and shares the same physical file as INST. The only difference is that INSTC uses different tags from INST. It is a part of CARD, and only supports the outputs of catalogues and indexes in various formats.

11. STAC, BLEV, DOTY, MEDI, LANC etc., RD (KSAM)

These are all the RD(KSAM) databases used for code validation check. For instance, STAC contains all the fields related to status codes used in the system; BLEV for codes of bibliographical levels; DOTY for types of document, etc.

The above databases are the main ones in the acquisition, cataloguing and public access subsystems.

Functions of the Subsystems

Acquisition subsystem

Acquisition staff collect library materials in different languages and different types through ways of purchasing, exchanging or gifts according to the acquisition policies. Once the material is decided upon, acquisition staff create a new record in the system, and are responsible for receiving the material, modifying the record, putting it into the account, printing out the cards, etc., and finally transfering the material to the cataloguing subsystem with its worksheet. According to this routine, the acquisition subsystem supports the following functions:

- Creating new records with worksheets.
- Duplicate checking with the assistance of QUERY.
- Printing out ordering cards, claim sheets, and account sheets, etc.
- Controlling acquisition records and worksheet outputs with acquisition status codes.
- Checking, and modifying records of received materials.
- Printing out worksheets for cataloguing.
- Managing data such as dates, prices, copies as well as bills, and supporting accounting.
- Supporting statistics based on subject areas, types of materials in order to control the acquisition quality.

Cataloguing subsystem

Cataloguing staff receive materials together with their worksheets from the acquisition. They fill the blanks in the worksheets, and then give the worksheets to professional data input clericals. Cataloguing staff fill another worksheet of analytical materials, and create records. The clericals modify acquisition records (sharing the same physical files) with cataloguing data; and complete the records with their printed catalogue cards. Functions of the cataloguing subsystems are:

- Modify acquisition records
- Create new records of analytical materials with worksheets.
- Support printing catalogue cards in batches.
- Print catalogue cards and other types of catalogues in standard formats such as AACR2, and the national standards
- Printing indexes and subject indexes in different formats.

The subsystem of public access will be used in 9 service desks in the library. The subsystem provides users with about 40 inverted fields, such as titles, authors, subject keys, etc. The Thesaurus of Chinese Mechanical Engineering will be attached to the subsystem, so as to allow online searches with the thesaurus assistance.

As for the details of functions of the subsystems, database structures, lessons and experiences, we will explain them later at convenience.

Finally, we would like to thank IDRC staff and librarians, as well as Mr. Wu Kang for their kind helps and advices.

A demonstration of the system will be given in the afternoon. We are looking forward to meeting you there.

机械中心图书馆计算机联合系统
中的采购、编目、检索子系统

机电部机械科技情报所
机械中心图书馆
李林甲　施明红

本文将简要介绍机械中心图书馆计算机应用的环境, 联合系统的设计原则及特色, 以及该系统中采购、编目、检索子系统的数据库结构及其功能.

机械中心图书馆概况

机电部情报所图书馆的正式名称叫机械中心图书馆, 它是由情报所的八个与图书馆业务紧密相关的处室组成. 机械中心图书馆也是我国机电领域收集、存储、检索和提供科技情报服务的国家中心.

该图书馆不仅向本所 800 余名专业技术干部提供服务, 而且有责任向全国范围的机电行业及其相关领域的所有单位和用户提供信息和服务. 馆藏 60 万册的书刊资料以及 6 种机读文献构成了该图书馆对外服务的信息资源基础.

图书馆拥有 2 套 HP3000 小型计算机, 带有约 50 台终端设备, 并拥有 20 台微型计算机. 图书馆日常工作中 MINISIS 的应用项目, 始于 HP3000／Ⅲ计算机. 目前, 图书馆计算机联合系统的项目, 通过两台支持中西文处理的 HP2392A 的仿真终端进行建库工作.

系统设计原则及特点

1987 年 10 月, 图书馆成立了自动化小组. 经过近一年的调研及系统分析, 拟定了《以 MINISIS 为基础开发图书馆计算机联合系统的总体规划》. 根据这一《规划》, 整个图书馆的工作将由采购、编目、报导、检索、流通及期刊管理六个计算机化的子系统来实现.

该系统的设计原则及其特点如下:

1. 建立两个 MINISIS 主文件

该图书馆的主要馆藏可大体分为三种: 图书、文献和期刊. 由于描述和处理期刊的方式与图书文献很不相同, 因此决定建立两个主文件系统, 一个用来处理图书文献, 另一个用来处理期刊.

2. 多语种处理系统

该图书馆的馆藏语种主要有汉语、英语、德语、法语、日语和俄语等. 具体的处理方法是: 将不同语种的馆藏综合在同一个主体文件中处理. 如: 图书和

文献无论是什么语种的，都在同一个数据库系统里来处理。每个字段也尽量设计成可处理不同语种的数据，除非这种不同语种数据共用一个字段的方式不能解决问题。如："出版年"，"出版地"，"文献类型"等字段都可以用来处理不同语种的数据。但是象题名、作者这样的字段，设置起来要格外小心。由于中文和西文题名字段的抽词方式大不相同，所以该系统将此类倒排方式不同的数据设置为两个字段来分别处理。

3.采用工作单录入数据

选择用工作单而不是用联机录人的工作方式有三个原因：

a.联机混合录入多语种数据很不容易，应该安排专职录人人员来做此项工作。

b.采用工作单的录入方式，有利于收集、分编工作人员集中精力搞好业务工作；录人人员发挥录入技术优势；并且有利于手工工作方式尽快转换成计算机辅助的工作方式。

c.进人数据库的数据可分别进行多次检查，进而提高了数据库的质量。

4.对一次和二次文献进行综合处理

该图书馆的主要馆藏是各文种文献，如专题论文、技术报告、会议文集等。大多数文献资料不但需要进行编目处理，而且还需对其进行深加工，并以文摘题录或其它的索引形式进行报导。因此，系统的开发基于对一、二次文献的综合处理。

5.建立必要的验证文件和权威文件

为了控制录人数据的质量，该系统采用了多个验证文件和权威文件。如状态代码验证文件、语种代码验证文件、还有机构代码权威文件，该文件用来在输人团体作者字段时使用代码，而不必录人全称。

6.采用书目级别控制

由于该系统综合处理一次和二次文献，采用书目级别控制可以分清每个记录的性质，同时也为编辑各种文摘或索引提供了方便。

7.支持标准化的输出格式

本系统的所有字段，在设计时就考虑到支持输出符合国家标准的或国际标准著录规则的卡片或索引。

8.采用词表

为了提高多文种馆藏数据库的检索能力，我们将利用《中国机械工程主题词表》建立与馆藏库相接的联机词库。该词库将装人汉字和英文主题词及与其一一对应的主题词代码和汉语拼音。由于 MINISIS 的 THLOADER 模块支持翻译功能，同一概念不同文种的资料无论是用汉字或英文的键词、以及键词代码或汉语拼音都能检索到。

采购、编目、检索子系统的数据库结构

为满足采购、编目、检索子系统的功能需要，该系统主要设置了以下数据

库:

1.LIBRARY RD MASTER 图书文献主数据库

LIBRARY 是这些子系统的核心数据库。它含图书馆采购、编目、检索工作所需几乎所有的字段。该数据库处理三类数据：日常工作管理用数据，书目描述数据，以及一些统计数据。

2.ACQ PS 收集数据库

ACQ 是 LIBRARY 的一个 PS 库,支持图书馆日常的收集工作,并含所有收集工作所需的字段。该库用来收集除期刊外的属于 A、M、C、S 四种书目级别的各类资料，并用于帐页的输出。

3.ORDER DS 订购库

该数据库含 ACQ 库的所有字段, 外加与书商代码有关的字段, 如书商全名、地址等; 另外还含与团体作者代码有关的字段。因此可以由此 DS 库输出有书商和团体作者全名的订购记录。此库仅供采购人员输出预订卡和催到单。

4.WKSHEET DS 工作单输出库

这是另一个收集工作人员专用的工作单库。它含分编人员工作所需的所有字段，供收集人员在到货验收并修改完采购记录以后，输出编目用的工作单。

5.CAT PS 分编数据库

CAT 也是 LIBRARY 的一个 PS 库。它含分编和标引工作所需的所有字段。在分编人员填完工作单上由收集人员留下的空白项目后，录人员通过此数据库修改数据，并完成馆藏记录。

6.CANAL PS 文献报导数据库

CANAL 是另一个 LIBRARY 的 PS 库。它仅用于录入由分编工作中产生的新记录。如: 大多数的分析级文献。这些用于报导的析出文献的整个记录是由分编人员，而不是采购人员生成的, 且与采购记录之间没有联系。

7. PUBLIC DS 检索数据库

PUBLIC 是用于查找图书馆馆藏的检索数据库。它含 CAT 和 CANAL 库的所有字段，外加与团体作者代码及全名有关的字段。该库设初始值限制, 用户不能访问正在处理过程中的记录, 也不能访问某些只为工作人员查看的记录。该库是为用户及服务台工作人员使用的数据库，并且支持各种格式检索结果的输出。

8.CARD DS 卡片输出库

CARD 库与 PUBLIC 库相似, 但未置初始值。此库供分编和标引工作人员核查新建的馆藏记录，并且支持输出不同格式的目录卡片、书本式目录和索引记录; 另外可供服务台工作人员查看 PUBLIC 库中读不到的记录。

9.INST RD (KSAM) 机构权威文件

INST 含描述机构或团体所需的所有字段, 如机构、团体代码, 全名, 地址等。它是团体作者及出版机构的权威文件，不能够单独使用。

10.INSTC RD (KSAM) 机构名称库

INSTC 是与 INST 定义几乎相同的数据库, 它不但含 INST 的所有字段且

与后者共享一个物理文件。其唯一的区别在于它们的字段标识不同。原因在于INSTC 作为 CARD 库的一部分，用于输出含有机构全名的各种格式的目录。

11.STAC, BLEV, DOTY, MEDI, LANC 等 RD (KSAM)　验证文件库

这些都是 RD (KSAM) 型数据库，用于各种代码录入时的验证。如：STAC 含所有与状态码有关的字段；BLEV 含与书目级别代码有关的字段；DOTY 含与文献类型代码有关的字段，等等。

上述这些数据库构成了采购、编目、报导、检索等各个子系统。下面分别介绍其功能．

采购、编目、检索子系统的功能

采购子系统

收集人员通过订购、交换、赠阅等方式，收集符合本馆收集原则的各文种、各类型资料。经查重确定入藏后，通过收集数据库建立新的采购或收集记录，并负责到货验收、修改记录、做帐、打卡、统计等工作。最终将资料与"编目工作单"一并转入分编部门。

该系统的主要功能有：

—以工作单方式建立采购记录。

—自动查重，并辅助 QUERY 查重。

—能够打印预定卡、催到单、帐页等。

—能够生成分编工作单。

—能够用状态码动态控制采购记录及工作单的输出。

—到货验收。对预订记录进行相应修改，使其与实物保持一致。

—对日期、价格、数量及发票等进行管理，定期打出帐页，并统计收支情况。

—能够按专业、按文献类型等统计馆藏，随时控制采购工作的质量，并生成统计表格。

编目及文献报导子系统

分编人员收到来自采购部门的新到资料及分编工作单后，对其进行分类、标引等进一步加工处理；同时对需作报导的析出文献，要另填工作单并通过报导数据库建立新记录。所有数据由专职录入员录入，分编人员负责核查并做有关输出。

该系统的主要功能有：

—能够以工作单方式对新到资料进行联机编目，产生新的馆藏记录。

—支持批处理的工作方式。

—能够输出符合"英美编目条例"（AACR2）及国家标准著录规则的馆藏目录卡片、书本式目录及各种专题目录、索引等。

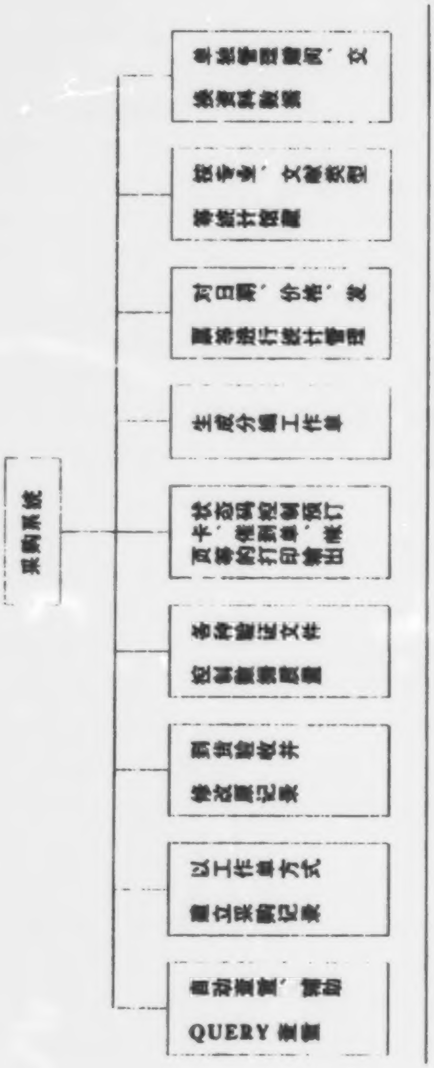
—生成符合排版要求的二次文献报导清样。

检索子系统

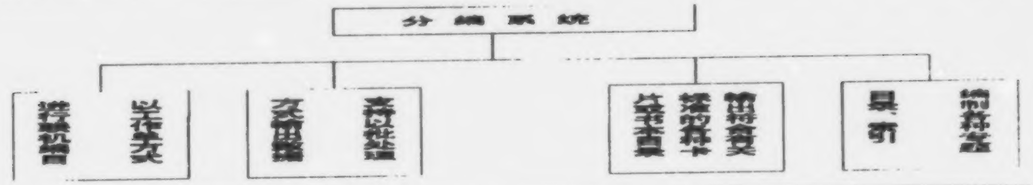
本系统主要由几个 DS 型数据库构成，将用于图书馆的九个阅览室．系统为用户提供约 40 个倒排字段，如题名、作者、分类、题内键词等，并支持各种逻辑检索提问式．待《中国机械工程主题词表》装入后，用户将能够用主题词联机查找馆藏文献．

各子系统的详细功能、数据库结构及设计过程中的经验与教训，我们将在演示中做出解释．

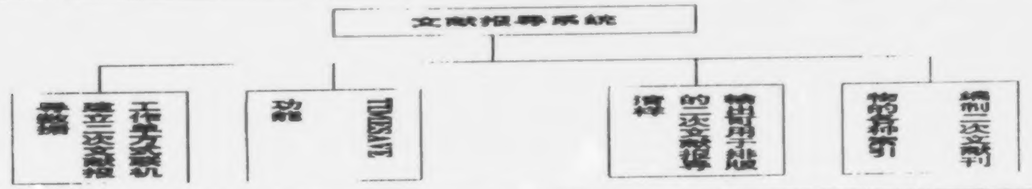
最后，对 IDRC 的朋友和馆员们及吴亢先生给予本项目的指导与帮助，我们表示衷心的感谢．


采购系统
单独管理赠阅、交换资料数据
按专业、文献类型等统计馆藏
对日期、价格、发票等进行统计管理
生成分编工作单
状态码控制预订卡、催到单、帐页等的打印输出
各种验证文件控制数据质量
到货验收并修改原记录
以工作单方式建立采购记录
自动查重、辅助 QUERY 查重

采 购 系 统 功 能 图



分 编 系 统 功 能 图


文献报导系统
TIMESAVE
功能
输出可用于排版的二次文献报导清样
编制二次文献刊物的各种索引

文 献 报 导 系 统 功 能 图

MINISIS IN THE CYRILLIC AND GREEK ALPHABETS AT THE MULTILINGUAL BIBLIOSERVICE OF THE NATIONAL LIBRARY OF CANADA

Virginia Ballance

ABSTRACT: The implementation of MINISIS at the Multilingual Biblioservice (MBS) has been phased in: first all roman-script language collections were automated and then the Cyrillic (Russian and Ukrainian language collections) and the Greek alphabet collections were automated. Ultimately the system will contain bibliographic data for 32 languages and will contain 10 different scripts. This paper will describe the MBS and review the progress of its automation project to date.

Although English and French are the official languages of Canada, and are the dominant cultural forces in Canada today, many Canadians come from other linguistic and ethnic backgrounds. In recognition of this, the Canadian government is committed to promoting and encouraging Canada's multicultural heritage. My department, the Multilingual Biblioservice, is the National Library's contribution to Canada's policy to promote multiculturalism. MBS was founded in 1973 to help the public library systems throughout Canada serve their multicultural or ethnic populations. The MBS lends books and books-on-cassette, such as novels, cook books and children's books to public libraries in 32 languages.

The idea to establish a centralized service was adopted because the costs of acquiring, processing, cataloguing and circulating books in foreign or unfamiliar languages was prohibitive for public libraries. In addition, much specialized knowledge and skill is required when selecting titles for purchase, in finding book dealers or suppliers and in knowing how to catalogue the material once it has been received by the library. And finally, if the group served is very small, the books purchased by the local library will be quickly read out and the readers will want more!

As a result, MBS is a central agency serving the whole country. We select the titles for purchase and acquire them from book dealers publishers or book stores either in the country of origin or from a local Canadian dealer. We purchase many copies of each title depending upon the demand for the language. For example we purchase 7 or 8 copies of every title in Chinese, whereas in Latvian or Welsh we buy only one copy. These books are then catalogued in the language and script of publication. We feel that transliteration or romanization of non-roman scripts is useful only for librarians, not for users. In addition, transliteration or romanization schemes vary, are not reversible and are often meaningless to native speakers. Instead of subject headings a brief contents summary note is written in English and French which describes the content of the book, for example: A historical novel. Roman historique. Once the books are processed, bound, labelled etc., they are circulated on long-term loan (from 1 to 10 years) among the ten provinces and 2 territories of Canada. Once the books in a particular province have been to all the libraries, they are returned to MBS where they are redistributed to another province.

From its beginnings 16 years ago, MBS has grown to a collection of 32 languages in 10 different scripts. The collection totals over 100,000 titles and 350,000 volumes. As well as books, we have a small collection of books-on-cassette (also called talking-books or audio-books) and books printed in large print.

Until 2 years ago all aspects of our operation were performed manually. This was especially difficult for the circulation function because we kept a record of the circulation history of every volume in the collection. Naturally automating our collection was imperative.

We chose MINISIS over all alternatives for one reason: its ability to handle data in an extended roman script and in other scripts. All the many wonderful features of MINISIS were simply a bonus. Even the National Library's system DOBIS was unable to meet our single requirement that the system handle data in other scripts.

Our use of MINISIS might be described as a fairly standard integrated library system. We have 3 rds: CATALOG for bibliographic and circulation data, ORDER for ordering information and ADDRESS for vendors names and addresses. These three data bases are linked in a ds called ACQUIS for the production of purchase orders. The CATALOG data base is MARC compatible meeting the Canadian minimal level standard.

The system is now fully implemented. Books ordered using MINISIS are now fully processed and are in circulation. MINISIS has helped us produce purchase order slips, labels and cards for processing the books, select books to send to a province and finally produce the catalogue to accompany the shipment of books to a province.

Where our application is not so typical is that the data may be in any one of 24 languages and 3 scripts.

What we have done is extend the basic roman or latin alphabet ASCII character set so that every available space in the set has been filled with either an accent/letter combination or special character. This character set is stored in a terminal handler which interprets the value and, by using special terminals which allow the user to define character sets, it may then be displayed correctly on the screen. To accommodate all the accented letters and special characters found in the roman script languages MBS has defined two character sets: one for all European roman script languages and one for Vietnamese. Vietnamese has its own character because it has so very many accents and special letters.

In order to access a particular character set, the user must first load the set into the memory of the terminal. The special characters and accented characters are stored in the upper level of the ASCII table, therefore, in order enter these letters, the user must shift to the extended table by means of a function key or entering an escape sequence.

The latest development at the MBS has been the implementation of the Cyrillic and Greek character sets on our system. With this development we are now able to automate all the titles in our Russian, Ukrainian and Greek collections. These character sets use the 2 bit alternate character set code that is built into MINISIS. As with the European character set, every letter in the two alphabets has been defined in the upper level of the ASCII table and the character sets are stored in the terminal.

When entering data the Cyrillic or Greek alphabet is preceded by an escape sequence or "shift in" (stored in a function key) which moves the user into the alternate character set. The character set goes through a terminal handler so that all terminal display will immediately be in the Cyrillic or Greek alphabet. When all the Cyril]ic or Greek data has been entered, the user must "shift out", to move back into the roman script, again by means of a function key. In this way, data in two scripts may be entered in the same field. This method of data entry works well as long as the user remembers to shift back into the roman script before hitting the enter key, otherwise all the MINISIS dialogue will be displayed in English but using Cyrillic or Greek letters until the next "shift out" is entered. It is possible to add a "shift out" at the end of field if necessary. To ease input, transparent key-cap overlays with the the Greek or Cyrillic alphabet were purchased to transform our keyboards from the roman alphabet to be bilingual Roman/Cyrillic or Roman/Greek keyboards. Letters were assigned their positions mnemonically, so Cyrillic A is on Roman A key and the Cyrillic C is on Roman S key, etc. The letters unique to the Cyrillic and Greek alphabets were placed on the top row, over the numbers, so it is necessary to shift out to enter any numerals.

Output of Cyrillic and Greek has proven to be a little more difficult than first anticipated. Strictly speaking dot matrix printers should be capable of producing Cyrillic and Greek letters however we have discovered that when a printer manufacturer claims that the product can print the Cyrillic alphabet, they really mean Russian. It will not necessarily print the special letters that are found in Ukrainian, Belorussian or any of the other languages that also use the Cyrillic alphabet. Our MINISIS distributor, McLeod-Bishop Systems will be working on this problem in the coming weeks.

With the implementation of the Cyrillic and Greek alphabets on our system, close to 75% of the collection is accessible on-line. The next phases are no less challenging as we must automate the Arabic script languages (Arabic, Urdu and Persian), Indic script languages (Hindi, Gujarati and Panjabi) and, of course, Chinese and Japanese.

Fortunately the significant developments by our fellow users in implementing these scripts on MINISIS will be very helpful to us at MBS in the future.

The Periodical Computer Management System based on Combination of Chinese and Western Languages (PCMS)

JIANG WEI FEZ
ZHAO HAS JUN
Airforce Political College Computer Centre
Jiang Wan Wu Jiao Chang
Shanghai China

KEYWORDS

Periodical Management; Database; Library automation.

ABSTRACT

This paper illustrates how to create a periodical computer management system (PCMS) using the MINISIS relational database management system on an HP3000 minicomputer system, and describes the system design consideration and functions of the database, and discusses the maintenance and safety of the database.

1. Prelude

Periodicals, an important kind of publication, has its own significant place in the modern literature field. Sixty to seventy percent of scientific information comes from periodicals. To give a better role to periodicals and the retrieval of periodicals, we have developed the periodical computer management system (PCMS). PCMS is divided into three main parts: the current year periodical management, non-current year periodical management and index to the title of periodicals. It has many functions. It can increase subscription renewal, stop subscription, check and accept, check-in and print check-in report of situation, and so on. It can also catalogue non-current year periodicals, perform multichannel retrieval, and retrieve by title in periodicals with subject words, generate statistics on indexing vocabulary, and so on. The software operates on the HP-3000/52 minicomputer system which was put into service in the library in the Shanghai Airforce Political College this year.

2. System Design and Function

PCMS uses the MINISIS relational database management system. MINISIS software is a system developed by International Development Research Centre of Canada (IDRC). It can deal with variable length field and subfield, either repeatable or non-existent. No space is needed for the non-existent field which saves disc space. All these characteristics make it suitable for management and retrieval of periodicals. That is why we have decided to develop a periodical computer management system using MINISIS.

The subscription for the periodicals averagely lasts for one year or longer. The increase in amount of data is slow relatively, unlike the rapid increase of the books. So it has relatively fewer data. It is more stable and has high rate of repeated use. An input of the periodical data could be used for one year generally, and usually, only the current number and current volume are added in every check-in. We divide the management of periodical into the current year periodical management and non-current year periodical management according to the demand in managing periodicals. The fields needed by the index to the title of the periodical are different from those in periodical management. For this reason and for the convenience of the application and maintenance of database, we create a separate database for the index to the title of the periodicals.

(1) Management of the current year periodicals
We have defined the field attribute of the current year periodicals to meet the demands for the periodical management, to design as many functions as possible, and especially to accomplish the work which is very hard or impossible with manual labour. There are 10 fields which include: periodical coded number (A100), the name of periodical (A200), publication period (A300), amount of subscription (A400), cost (A500), check-in (A600), location (A700), location number (A800), signal for stop of subscription (A900) and the place of publication (B100). The A100 and A200 in the parenthesis indicate the tag of the field. The operation and application that follow next will take the field tag as the operation target. Periodical only has one periodical coded number and it is printed in them. If there are two same periodical coded numbers, the system will automatically warns:

THIS MAY BE A DUPLICATE RECORD
POTENTIAL DUPLICATES. DISPLAY (Y/N/ALL)?

Input Y, it will output the whole record. We also define the periodical coded number as inverted item and build up a inverted file, so as to find the record of periodical quickly. For functions such as subscription and check-in, the operator can find the record by the periodical coded number. For those periodicals which are not published by post office, periodical management can assign a coded number according to a certain rule and the circumstance of the library, thus treating these periodicals as the periodical published by the post office.

The name of periodical and location number are both defined as inverted items. Thus users can find a certain periodical quickly by referring to A200 and A800. The multichannel retrievals are therefore possible. Location number is defined with a length of 6 bytes: the first is the classification number, indicating that it belongs to that kind of periodical. The classification number is divided into 22 major classes, for example, N indicates natural science, and B indicates Philosophy. This make it possible to count up the subscription amount for various periodicals and expenses.

It can help the subscribers to make decision so that they can control the amount of the periodical subscription.

The Check-in field has two subfields: current number (A601) and the amount of receipt (A602). The amount of receipt means the actual amount of the periodicals received. The Check-in field is defined as a repeatable field, which can repeat up to 200 times. This is enough for every check-in. Most periodicals are checked-in once a month and some are checked-in half-monthly or bi-monthly. To print check-in report table, we have also defined the current number (A601) as an inverted item to quickly find whether a record has been checked-in or not.

Location (A700) has six subfields which include: students (A701), teachers (A702), foreign (A703), respectively the reading room for students, reading room for teachers, foreign books reading room, restricted materials reading room and periodicals storage.

We have also designed the field of stop subscription. For those periodicals which will not continue subscription, input "s" in the field. We define it which does not display the name of field automatically on the screen to increase the speed of input. Since there are periodicals with the same name, we added a publisher field. The field, when input, does not display automatically. It does not input data for this field until it has discovered the periodicals of the same name. There are not many periodicals of the same name. We can find several such periodicals in the inverted file.

We have designed two major functions of subscription and check-in in the management of the current year periodical.

(-). Subscription
The function of subscription includes continuing, stopping and increasing subscription. The function to continue subscription searches for the periodical coded number. If it finds one, it will treat it as periodical to be continued, or it will treat it as periodical to be enlarged in subscription. After transferring the record onto the terminal screen and checking the contents of the record, we can make relevant modifications, for example, the price will probably change every year. With a purge of data in the check-in field, the subscription is discontinued. As for function to stop subscription, you only need to input letter "s" into A900 field and purge data in the check-in field. The functions to continue subscription and stop subscription are done by choosing the MINISIS MODIFY processor. To increase a subscription, we input the record for the newly subscribed periodical by choosing MINISIS ENTRY processor. The records of the subscribed periodical, after they have all been input into computer, can be printed on a subscription order and the total price can be calculated.

(=) Check-in
After you have received the periodicals, you can find the record by the periodical coded number. If there is no periodical coded number, the periodical is not subscribed to by this library, otherwise, you just go on with your work. You input relevant data into the two subfields that is the current number and the amount received. After that, check-in is finished. After all the periodicals received have been checked-in, you can find the records (both registered and un-registered) with the MINISIS QUERY processor and print a check-in report of situation with the MINISIS PRINT processor.

(2) Management of non-current year periodical
Decided by the management of the non-current year periodical, we define the attribute of the field in the record. The record has 11 fields, which are ask for number (B500), periodical coded number (B700), the name of periodical (A100), responsibility area (A200), volume issue number year month (A300), publish, (A400), media (A500), periodic publication (A600), explanation (A700), ISSN (A800), collection (A900). To make it suitable for various periodical, the name of periodicals (A100) has three subfield: the main name of periodical (A101), the parallel name of periodical (A102), vice-name of periodical (A103). The parallel name of periodical is defined as a length of 120 byte, to be used to input the English name.

A900 has five subfields: registration number (A905), year of publication (A901), the amount of collection (A902), lacking number (A903) and the place of the storage (A904). Because the periodical coded number and the ask for number are one, we define it as the field which is checked for duplicates. It also an inverted field. To make it possible to search quickly, A101 and A200 are defined as inverted fields.

The amount of data of the non-current year periodical is much bigger that of the current year periodical. The historical process of the periodical publication is recorded. Considering this factor, an explanation field is defined as a repeatable field with a length of 100 bytes to be long enough to store the data. These fields are variable in length and will not waste space.

Usually, the current periodical, after one year, will be bound and made available for the readers. Readers still can search cards manually, but they can also search using the computer to do this with the B500, B700, A101, A200 fields.

(3) Index to the title of the periodicals
To make the index to the contents of the periodicals in every month and to provide readers with useful and timely information, we have developed an index system of the title of the periodicals.

The index system is divided into three databases: Index DB, code DB and title DB. These DB are respectively called INDEX, KSIN and INDS. INDEX and KSIN are both Master RD (relational definition) DBs. INDS DB is DS (Data Submodel). INDEX and KSIN are joined by the code field, forming the INDS DB.

The reasons for designing three DB are: many different titles of articles may come from the same periodical. We use the code to replace the periodical to as to reduce the repeatable input thus saving both space and time.

There are five fields in INDEX DB which are: title of the article (A100), author (A200), subject word (A300), the code of the periodical (A400) and issue number (A500). A100m A200 and A300 are all inverted fields. A200 and A300 are also defined as repeatable fields.

There are three fields in the KSIN DB: code of the periodical (B200), the name of the periodical (B100) and publisher (B300). INDS DB is the combination of the INDEX and KSIN, so its fields are the total sum of the fields in INDEX and KSIN.

The workers input the relevant code of periodical into KSIN and then input the data field into every field of INDEX. Users can retrieve by title using the subject word in the INDS DB and then print the index entry of the title.

By using the inverted file, you can also generate statistics on the indexing vocabulary.

(4) Design various print formats
To satisfy various speciality demands of the management and retrieval of the periodical and make the work easier for managers and readers, we have made use of the MINISIS PRINT processor and have designed five print formats which are: print subscription order, print non-current year periodical catalog card, print title index, print check-in report of situation and print not check-in report of situation.

3. The maintenance and safety of database

The set up of DB is the initial part of the work. Maintenance is needed after the set up DB to safeguard its normal operation and to improve function. MINISIS DBMS provides a series of utility processor, helping DB manager to expand the physical store space and invert file space and reclaim lost space and logical reconstruction and store/restore and recovery lost data. We created a logfile. If failure happens in the system, we can recovery the DB using the backup and logfile.

To protect the DB from being used by illegal users, we have taken a series of steps to guard the data. HP3000 has strict and fine technology to detect the identification of user. Data is not available for illegal users. Each user has a password. There kinds of user are: QUERY (searcher), INPUT (the person who enters data) and DBA (database manager). By making use of the SECURE command of DATADEF processor of MINISIS, SECURE gives the Database Manager the ability to limit user access to both databases and programs by creating security profile for each user in an account, the user name of QUERY and INPUT and DBA have been stored in the user's security profile. QUERY users can only execute the QUERY and PRINT processors, INPUT users can only execute the ENTRY, MODIFY, PRINT and QUERY processors, and DBA user can execute all the MINISIS processors.

4. Brief summary

At the very beginning of its service of this system, the work to input data is heavier. It will take two skilled input workers about two weeks time to do this work if we suppose that library has subscribed 2000 kinds of periodicals. When the input is finished, we can use these data repeatedly. The amount of record increase little every year, with most of them just adding or deleting certain fields . This is different from the book management system which always keep increasing in large quantity.

The check-in of the periodical was formerly done by hands marking on a card which was quite inconvenient, especially when calculating the total price and printing subscription orders and check-in reports of, and so on. After adopting this system, it becomes convenient, accurate and timely.

The index to the title of new periodical provides readers with timely information.

In a word, the realization of the system greatly improves the efficiency of work, reduces labour, and make thing easier for managers and readers. It realizes automation of the periodicals management.

Persons involved in the development of this system have also Zheng Chen Yang and Lin Ping Zhong. Congratulations on our successful co-operation. Thank you!

Warnier Methodology Applied to the Integral Design of a Library System Using MINISIS

M.en C. Ivonne Guzmán
Lic. Ofelia Solis
Lic. Ricardo González
Universidad Autónoma Metropolitana
Mtra Luz Marina Quiroga
Mtro Roberto Garduño
Universidad Nacional Autónoma de Mexico

Abstract

This paper presents the background and the principal points considered in the analysis and objectives of the first stage of the project. Also described is the application of the Warner methodology to the design of the logical data bases using a concrete example, and how the physical MINISIS data bases were derived from the logical data bases.

Background

The U.A.M. is composed of three campuses located in Mexico City. Each campus is independent of the others and has its own infrastructure, which includes its own library and computer center. The book collection, classification and control of library materials are done according to the needs of each campus.

The development of the system will follow two phases. The first incorporates three modules: selection and acquisitions, bibliographic analysis, and circulation. The second incorporates the following modules: Serials control y audiovisual materials. The system works on the basis of distributed processsing and distributed data bases, allowing each campus library to assign the topographic number most appropriate to the academic needs. The HP 3000 model 52 of each campus will be interconnected by means of the UAM network so that searches can be made in any campus data base by author, title and descriptor.

Since November 1988, work has been done to standardize cataloguing and to produce a cataloguing worksheet and standards manual.

In general terms, the following were defined in the analysis of the information requirements:

- the areas that were using the most human resources and had the highest flows of information;
- operations of the actual system;
- inter-relationships with other applications, specifically the relationship with the student control application; and
- the impact of the system on the university.

Objectives of the System

The first stage of the system will have the following:

1. For the Bibliographic Analysis module, the realization of:
- Catalogue cards by author
- Bibliographies by subject
- Cataloguing statistics for productivity analysis
- Labels required for processing (circulation cards, etc.)
- On-line cataloguing

2. For the Acquisitions and Selection module, control of:
- Number of titles aquired monthly
- Titles passed on for technical processing
- Number of books bought per collection
- Titles ordered by not yet received
- Purchases by department
- New titles received quarterly
- Books ordered monthly
- Books donated each month
- Number of copies of a book available for exchange with other libraries
- Titles received quarterly from other libraries
- Number of copies of a particular title
- Total budget available, spent and remaining
- Purchase orders not yet paid
- Status of the orders
- Price information on each book

3. For the circulation module, control of:
Control of overdues

- Monthly statistics of books in external circulation by collection
- Usage statistics of the titles in internal circulation
- Lists of damaged books to be replaced
- Books ordered by acquisitions to replace damaged books
- Interlibrary loans
- Overdue notices to staff
- Printing the date due slips
- Movement of copies between collections
- Title and copy being lent
- Return of a title/copy
- Retrieval of a particular copy from the data base by acquisition number
- Retrieval of books borrowed by a user
- Reserving books
- Retrieval of copies by author, title or call number
- Control of external users permitted to borrow books
- Payment of fines

System data base design

The methodology of Jean-Dominique Warnier was chosen for the system design, in light of the emphasis in the MINISIS course placed on defining system outputs, so that one could come up with the fields required in the data bases and their corresponding inversion characteristics.

The Warnier methodology focuses on the study of data structure, and in particular on the outputs of the system to arrive at the final physical data base structure, system inputs and system processes.

There are two fundamental concepts:

- the products or outputs of the system are the starting point for the system design
- the structure of the data determines the structure of the process

Described below, step by step and using an example, is the procedure used to define the logical data base. Described later is how the physical MINISIS data bases were derived from the logical data base.

System output definitions

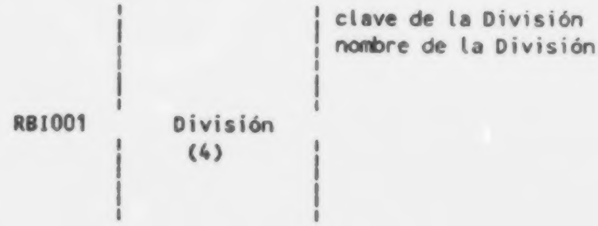
Appendix 2 gives an example of an output from the system. The first page gives general information about the report, indicating which fields have been sorted as defined by the user. The second page gives detailed information of each field found in the report, such as the name of the field, the length and type of field; the column "foriegn" indicates if the field contains source or derived data along with a description of the field and how the data is derived for the field in the case that it has been derived from a particular procedure. At this point we have a layout of the report or output.

Beginning with the pages of the data definitions in the reports, one sees the logical data dictionary of the system to which a field for synonyms has been added. This field is used to identify possible inconsistencies in nomenclature. The length of the field is the largest defined in the reports.

The fourth page of Appendix 1 shows the distribution of information. Each field is identified by a number. On the previous page one finds explanations which correspond to each number. Each report is designed together with the user, to produce a complete, understandable representation of the data.

Producing the Logical Structure of the System Outputs

In the report "Reclamacion de Aduedo de Materiales" of Appendix 1, the information is presented by division. Thus "division" becomes a key in the Warnier diagram. Next, the attributes of the entity division are listed.



The number of possible values that one can have are put into parenthesis. When the value is large or undefined, one or two letters are placed in the parenthesis. In each division, there are users who have borrowed books. This is represented in the central part of the Warnier diagram.

```
        |            | clave de la División
        |            | nombre de la División
        |            |
        |            |            | clave del usuario
        |            |            | nombre del usuario
        |            |            |
        |            |            |              |nombre del título
        |            |            |              |que adeuda
RBI001  | División      usuario    Libros prest
           (4)       |    (U)     |    (LP)      |fecha de vencimiento
        |            |            |              |del préstamo
        |            |            |
        |            |            |
        |            |
```

As one can see, the structure of the report is determining the structure of the data. Information that can be derived from calculations from other data and constants should not be stored in the data base. In this example, the report date is obtained from the computer, the page number is a sequential number calculated by a program. Thus this data, as well as the page headings and messages, are not part of the logical design of the output. The number of books owed can be calculated by counting the number of books overdue, using the date. Hence, it is important to have the due date.

The same procedure is followed for all the other reports and on-line outputs defined by the user.

Producing the Logical Data Base

Arriving at the logical data base consists of eliminating unnecessary information, which is done in three steps: elimination of constant elements; elimination of elements that can be calculated from other data; and elimination of redundant elements.

Using the descriptions of the logical structures of the outputs, from which constant or calculated elements have been eliminated, one can proceed to defining the logical data base for each module. To illustrate how a logical data base is obtained for a specific module, the logical structure of the output defined previously will be used. From that structure, one obtains the following entities:

```
            | Clave de la División
Division      Nombre de la División
            |
            |


                      |
                      | Clave del usuario(matricula)
Ejemplares prestados    Vector de
                      |   Títulos en préstamo
                      |   Fecha de vencimiento del préstamo


             |
             |  Clave del usuario
Usuario         Nombre del usuario
             |  Clave de la División
             |
```

A user is part of a division, sothe division key is added to establish the relationship.

Whenever the attributes repeat, only one repetition is kept. One proceeds through all the attributes which come from the logical structures of the outputs. Finally, one obtains the logical data base for the module. This procedure is repeated for each module. In this manner, modules can be developed independently. To get to the integrated data base of the system, one repeats the procedures mentioned above using the entities of the modules defined. For example, consider the following entities:

```
                        | Clave del título
                        | Nombre del título propiamente dicho
Títulos en acervo       | Autores
B.D. del módulo         | Precio unitario
Selección y adqui_      | Fecha de compra
siciones.               |
```

```
                        | Clave del título
                        | Nombre del título propiamente dicho
                        | Editorial
                        | Precio unitario
                        | Fecha de compra
Títulos en acervo       | Letras de clasificación
B.D. del módulo         | Números de clasificación
Biblioteca              | Año de constitución
                        | Cutter de autor
                        | Cutter de título
                        | Año de publicación
```

```
                        | Clave del título (ISN)
                        | Título propiamente dicho
                        | Título paralelo
                        | Títulos alternativos
                        | Letras de clasificación
                        | Números de clasificación
                        | Cutter de autor
                        | Cutter de título
                        | Año de constitución
                        | Otra información sobre el título
Títulos en acervo       | Mencion de responsabilidad de la edición
de la B.D. Análisis       Menciones de responsabilidad subsiguientes
Bibliografico           | Relacionadas con edición
                        | Lugar de publicación
                        | Editorial
                        | Año de publicación
                        | Descripción física
                        | Material ilustrativo
                        | Serie
                        | Notas
                        | ISBN
                        | ISSN
                        | Clave-autor
                        | Clave-tomo
                        | Clave catalogador
```

From this example, it can be seen how the methodology allows one to define an entity that contains the attributes required by the three sections (acquisitions, technical processing and circulation.) It is also easy to identify what information is common to each section. This can be used in the design of a PS.

With this methodology, one avoids having redundant information even when each module is developed independant of the others.

Following through with the design of the previous entitities, the entity for storing information about books will have the following data:

```
                   | Clave del título (ISN)                         - o *
                   | Título propiamente dicho                       - o *
                   | Título paralelo                                    *
                   | Tiitulos alternativos                               *
                   | Letras de clasificación                          o *
                   | Números de clasificación                         o *
                   | Cutter de autor                                  o *
                   | Cutter de título                                 o *
                   | Año de constitución                              o *
                   | Otra información sobre el título                   *
                   | Area cronológica                                   *
Titulos en acervo  | Mención de responsabilidad de la edición           *
                     Menciones de responsabilidad subsiguientes*
                   | relacionadas con edición
                   | Edición                                          o *
                   | Lugar de publicación                               *
                   | Editorial                                        o *
                   | Año de publicación                                 *
                   | Descripcion física                                 *
                   | Material ilustrativo                               *
                   | Serie                                              *
                   | Notas                                              *
                   | ISBN                                               *
                   | ISSN                                               *
                   | Clave-autor                                    -   *
                   | Clave-tema                                         *
                   | Clave empleado                                     *
                   | Precio unitario                                - o
                   | Fecha de compra                                - o
```

The attributes marked with "-" will be utilized by the acquisitions and selection section. Those marked by "o" will be used by the circulation section and those marked by "*" by the technical process section.

The same process is done for each entity of the logical data bases of the different modules to be integrated into the system. There will be cases where the information in one entity will not be used by all the modules of the system.

One can implement each module separately as long as the physical files used have the structure of the integrated logical data base. This gives an integrated desing in the conceptualization of the system and modularity in the development and implementation of the system.

Once one has the complete logical design of the data base of the system, one proceeds with a retrospective analysis which should consider the following:

1. Make sure that each and every report and screen can be generated by the system.
2. If there is information missing, it should be added to the corresponding entity.
3. Identify those cases that involve the sequential opening of more than three files to get the information required. These cases may slow down the system. One can evaluate the possibility of including a redundant attribute in the system to increase response time. In the case of MINISIS, this may be done via a DS.
4. The logical data is normalized.

Design of the Physical Data Base

Using the normalized logical data base structure, the physical structure can be derived which depends on the hardware and the DBMS to be used.

If MINISIS is used, it should be decided what types of files are to be used, whether they will be MASTER or KSAM. The PS and DS should be identified. As an example, consider the following entities: Titles Catalogued, Subject, Authors and Copies. Looking at the information of Titles Catalogued, it can be seen that the data is of variable length and can be repeatable. This suggests that this entity be declared as an RD MASTER. We considered substituting the full author name and subject for the codes as had been defined in the entity title of the logical design. However, we left it as is (that is as codes) so that we could produce authority lists.

There can be various copies of a title, which suggests that an entity called "copies" be included as a repeatable subfielded field. However, the dat in the copies field is data that is subject to a lot of modification, which implies that GARBAGE would have to be run frequently. On a large data base, such as Titles Catalogued, this would take a great deal of time. So it was decided that the entity "copies" would be a data base MASTER RD independent of titles. Copies was then joined to Titles Catalogued by means of a DS using the title key and acquisition number, as was suggested by our logical design.

Conclusion

The participation of the users (library personnel) in the analysis is on-going. The users now know what to expect of the system and what options are available. This reduces the possibility that the system not be accepted because the users are kept up to date and involved in the design stages.

The Warnier methodology, based on outputs, allows for a systematic design that incorporates all the elements required in the physical data bases. The methodology avoids unnecessary information redundancies and allows for an integrated design while permitting an independent design and development of each module of the system.

From the logical data base designed using the Warnier methodology, the MINISIS data bases (contents and whether they are RD, PS or DS) can be deduced directly.

The system that we have designed can be utilized by other educational institution libraries with similar needs. Users of other libraries can evaluate each module (using the outputs as the base) and can then use some or all of the modules in their own system design. If such is the case, the staff of their respective computer departments can easily see if the modifications to the existing reports or if new reports can be generated directly from MINISIS or if it is necessary to make changes to the data definitions.

Appendices follow the Spanish version of this paper.

La Metodología de Warnier aplicada al Diseño Integral de Sistemas de Información Bibliotecarios usando Minisis

M.en C. Ivonne Guzman
Lic. Ofelia Solis
Lic Ricardo González
Universidad Autónoma MetropolitanaMtra luz Marina Quiroga
Mtro Roberto Garduño
Universidad Nacional Autónoma de México

Resumen

En este documento se presentan en forma sucinta los antecedentes y los principales puntos considerados en el análisis, asi como el alcance del sistema en la primera etapa. Se describe asi mismo la aplicación de la metodología para diseñar la Base de Datos lógica a traves de un ejemplo y las consideraciones para obtener la Base de Datos física en Minisis.

Antecedentes.

El análisis de las necesidades de información se inicio en febrero de 1988 y se obtuvo la primera aproximación de la Base de Datos lógica del Sistema para la Unidad Azcapotzalco, en octubre del mismo año; en este análisis se observó independencia en el manejo de información de las secciones (modulos) Hemeroteca y Material audiovisual, en relación a la información de los modulos Selección y Adquisiciones, Análisis Bibliográfico y Biblioteca (circulación).

En noviembre de 1988, las autoridades de la Universidad Autónoma Metropolitana U.A.M. decidíeron que se diseñara un sistema de información bibliotecario para la Universidad en su conjunto. La U.A.M. esta integrada por tres unidades (campus) distribuidas en la Ciudad de Mexico. Cada unidad es independiente y cuenta con su propia infraestructura, tienen una biblioteca y centro de cómputo. El acervo, la clasificación y control difieren en función de las necesidades de cada unidad.

El desarrollo del sistema se llevara a cabo en dos fases. La primera contempla los modulos: selección y adquisiciones, análisis bibliográfico y biblioteca. La segunda los modulos: Hemeroteca y Material audiovisual. El sistema operara con procesamiento distribuido y Bases de Datos distribuidas, permitiendo que cada Unidad asigne la signatura topográfica mas adecuada a las necesidades determinadas por los distintos departamentos académicos. El equipo HP-3000/52 de cada Unidad estará interconectado a traves de la Red-UAM, para poder consultar los catálogos de las otras unidades por autor, título y tema. Dadas las características del sistema, el diseño podrá ser utilizado de base por otras instituciones educativas en México o América Latina con características similares.

A partir de noviembre de 1988 a la fecha, se trabajo en la normalización de la forma de catálogar, en la elaboración de un manual de catalogación, así como en el diseño de una hoja de catalogación, en el análisis de los procedimientos actuales y en la elaboración de los nuevos procedimientos.

En lo referente al análisis de los requerimientos de apoyo informático se consideró lo siguiente:

Se identificaron las areas sustantivas que requieren mayor apoyo informático en función de los recursos humanos utilizados y volumenes de movimientos de información.

En relación a la operatividad del sistema actual se consideraron:

- Los objetivos de la biblioteca e interrelación de estos con los de la U.A.M.
- Los vinculos existentes entre las tres bibliotecas.
- Presupuesto en los tres ultimos años, distribución y crecimiento del acervo.
- Se revisaron los procedimientos, en cada caso se evaluó ¿Cómo se lleva a cabo?, ¿Por qué se hace?, ¿Quién lo hace?, el numero de pasos que involucra, la frecuencia; así como las deficiencias de operación actuales.
- Se identificaron las areas en que hay acumulación de trabajo y se analizó en que forma el sistema apoyará para disminuirlo.
- Se evaluó la suficiencia, disponibilidad, acceso y confíabilidad de la información y la dificultad para emitir informes. Se revisaron los reportes generados actualmente y los que se requieren, pero que son difíciles de obtener en este momento.

Interrelación con otras aplicaciones. Se detectó el vínculo que tendra el sistema con el de control escolar.

Impacto en la organización. La automatización facilitará el acceso y manejo de la información, apoyara en una mejor planeación y administración de los recursos, los procedimientos serán afectados, se contará con información completa, confiable y oportuna.

Alcance del sistema.

La primera etapa del sistema contempla:

1. Para el módulo Análisis Bibliográfico la generación de:

- Juego de tarjetas por autor
- Boletín Informativo (índice temático)
- Estadísticas de productividad en la catalogación
- Etiquetas para proceso físico (esquineras y tarjetas de prestamo)
- Catalogación en Línea

2. El módulo Selección y Adquisiciones permitirá tener control de:

- Ejemplares adquiridos en un mes
- Títulos que ingresaron a la sección de Análisis Bibliográfico
- Ejemplares comprados por colección
- Pedidos no surtidos
- Identificación de libros adquiridos para docencia o investigación
- Títulos nuevos que ingresaron a la colección en un trimestre
- Libros solicitados por los usuarios durante el mes
- Donaciones recibidas por mes
- Ejemplares disponibles para canje

- Títulos recibidos en canje durante un trimestre
- Nulmero de ejemplares por título
- Presupuesto comprometido, ejercido y disponible
- Facturas no pagadas
- Estado del pedido
- Datos económicos de un libro, por numero de adquisición

3. El módulo de Biblioteca (circulación) permitira tener control de:

- Reclamación de adeudos de material (alumnos)
- Estadística mensual de préstamo externo por colección
- Estadística de uso de títulos en préstamo interno
- Listado de ejemplares mutilados sujetos a reposición
- Libros solicitados a la sección de adquisiciones para reposición
- Vencimiento mensual de convenios de préstamo interbibliotecarios
- Generación de reclamos de adeudos a empleados
- Generación de la papeleta de devolución
- Ejemplares que cambian de colección
- Prestamo de ejemplares
- Devolución de ejemplares
- Consulta de ejemplares por número de adquisición
- Consulta de prestamos por usuario
- Apartado de ejemplares
- Consulta de ejemplares por autor, título o signatura topográfica
- Alta de usuarios al servicio de préstamo externo
- Pago de multas

Diseño de la Base de Datos del Sistema

Para su desarrollo se escogió la metodología de Jean-Dominique Warnier, dado el énfasis que se dió en el curso básico de Minisis a la importancia de definir el tipo de reportes y consultas que se requieren, con la finalidad de identificar que campos es necesario indexar y como organizar la información para poder realizar las consultas en forma expedita.

La Metodología de Warnier se concentra en el estudio de la estructura de los datos, y e particular de los productos de un sistema. Derivando de dicha estructura tanto las entradas al sistema como su Base de Datos y los procesos.

Son dos las ideas fundamentales:

- Los productos o salidas de un sistema, son el punto de partida para el diseño del mismo.
- La estructura de los datos determina la estructura del proceso.

A continuación se describirá paso a paso con un pequeño ejemplo el procedimiento para obtener la Base de Datos Lógica. Y posteriomete se enunciarán los parametros considerados para adaptar la Base de Datos Lógica al diseño físico de la Base de Datos (B.D.) usando Minisis.

Definición de las salidas del Sistema.

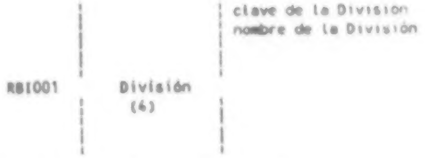
En el anexo 2 se muestra un ejemplo de la definición de un reporte o salida del sistema. En la primera hoja se presenta información general del reporte y se indica por que campos (datos) se requiere ordenamiento o selección de información, de acuerdo a las especificaciones del usuario. En la segunda hoja vemos una definición detallada de cada dato que se encuentra en el reporte, en la cual aparece el nombre del dato, el tamaño y tipo de dato; en la columna "formato" se indica si el dato es fuente o derivado, finalmente una descripción del dato y forma de obtenerlo en caso de ser derivado de algun proceso.

Posteriomente, a partir de la información contenida en las hojas de definición de datos de los reportes, se obtiene el diccionario de datos lógico del Sistema, al cual se le agrega un campo de sinónimos. Este campo sirve para identificar posibles inconsistencias en la nomenclatura. El tamaño de los datos será el más grande que se halla definido en los reportes.

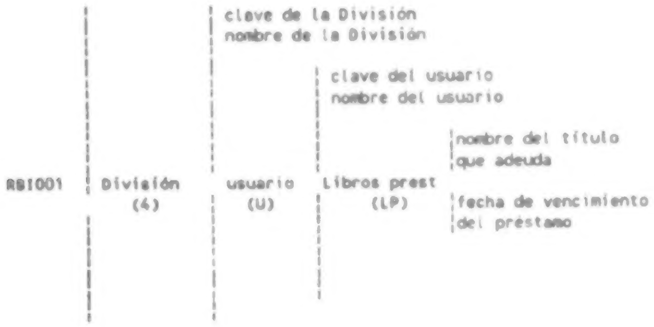
En la hoja 4 del anexo 1 se muestra la distribución qráfica de la información, identificando a cada dato con un número que lo relacíona con la descripción de la hoja anterior. Cada reporte es diseñado conjuntamente con el usuario, considerando la selección y ordenamiento de los datos que faciliten la visualización de la información de acuerdo a sus necesidades.

Derivación de la estructura lóqica de las salidas

En el reporte "Reclamación de adeudo de materiales" del anexo 1 observamos que la información se selecciona por División, por lo que en primer término el reporte RBIOO1 muestra selección de información por Divisiones. En la parte superior de la llave se indican los atributos de la División.


clave de la Division
nombre de la División
RBI001
División
(4)

Entre paréntesis se indica el número de posible valores que puede tomar, cuando es grande o indefinido se coloca una o dos letras de la Entidad. En cada División se prestaron libros a usuarios adscritos a estas. Esto se indica en la parte central de la llave.


clave de la División
nombre de la División
clave del usuario
nombre del usuario
nombre del título
que adeuda
RBI001
División
(4)
usuario
(U)
Libros prest
(LP)
fecha de vencimiento
del préstamo

Como se puede observar, la estructura del reporte va definiendo la estructura de los datos. La información que puede ser derivada mediante procesos de cálculo de otros datos o del sistema, no debe ser alamacenada en la base de datos, así como los datos constantes. En el presente ejemplo, la fecha de emisión del reclamo, se obtendrá del sistema de la computadora, el número de página del reporte y el número de ficha de reclamación son derivados de contadores por programa, por lo que esta información y los títulos y mensajes no son considerados en la estructura lógica de la salida. El número de ejemplares que adeuda se puede obtener de contar el número de ejemplares con fecha vencida no devueltos; es importante tener la fecha de vencimiento del préstamo.

El mismo procedimiento se sigue para todos y cada uno de los reportes o consultas por pantalla definidos por el usuario.

Derivación de la Base de Datos Lógica

El proceso de derivación consiste en la eliminación de información innecesaria, la cual se lleva a cabo en tres pasos:

- Eliminación de elementos constantes
- Eliminación de elememtos derivados
- Eliminación de elementos redundantes

A partir de la descripción de la estructura lógica de las salidas, de las cuales se han eliminado datos constantes o derivados, se procede a la derivación de la Base de Datos lógica de cada módulo. Para ilustrar el procedimiento de la derivación de la Base de datos de un módulo, considérese la estructura lógica del reporte RBI001. De esta estructura se obtienen las siguientes entidades:

```
            | Clave de la División
Division      Nombre de la División
            |


                       |
                       | Clave del usuario(matricula)
Ejemplares prestados     Vector de
                       |   Titulos en préstamo
                       |   Fecha de vencimiento del préstamo


             |
             | Clave del usuario
Usuario        Nombre del usuario
             | Clave de la División
             |
```

Un usuario esta adscrito a una División, por lo que agregamos la clave de la División para relacionarlo.

Cada vez que aparezcan datos repetidos para una entidad se eliminan y se van colocando los datos nuevos, que se deriven de las estructuras lógicas de los distintos reportes. Al final, se obtiene la Base de Datos del módulo y se repite el procedimiento para los diferentes modulos. De esta forma, se puede implementar independientemente cada módulo. Para obtener la Base de Datos integral del Sistema, se repite el procedimiento de derivación tomando como base las entidades de los modulos, por ejemplo:

Consideremos las siguientes entidades

Títulos en acervo B.D. del módulo Selección y adquisiciones.	Clave del título Nombre del título propiamente dicho Autores Precio unitario Fecha de compra

Títulos en acervo B.D. del módulo Biblioteca	Clave del título Nombre del título propiamente dicho Editorial Precio unitario Fecha de compra Letras de clasificación Números de clasificación Año de constitución Cutter de autor Cutter de título Año de publicación

Títulos en acervo de la B.D. Análisis Bibliografico	Clave del título (ISN) Título propiamente dicho Título paralelo Títulos alternativos Letras de clasificación Números de clasificación Cutter de autor Cutter de título Año de constitución Otra información sobre el título Mencion de responsabilidad de la edición Menciones de responsabilidad subsiguientes Relacionadas con edición Lugar de publicación Editorial Año de publicación Descripción física Material ilustrativo Serie Notas ISBN ISSN Clave-autor Clave-tomo Clave catalogador

En este ejemplo es posible observar, como la metodología nos permite definir una Entidad que contenga los atributos requeridos por las tres secciones y a la vez identificar que información es utilizada por cada una de estas para realizar las proyecciones (PS) correspondientes. De esta forma, el sistema ofrecerá a los distintos usuarios la visión de la información que les interesa. Con esta metodología evitamos tener información redundante al implementar los modulos por separado y solo se tendra la estrictamente necesaria para facilitar procesos o disminuir el mantenimiento a la B.D. Las posibles redundancias serán función de un análisis cuidadoso.

Realizando el proceso de derivación en las entidades anteriores, el archivo Títulos en Acervo de la Base de Datos del Sistema contendrá la siguiente información:

```
                          | Clave del título (ISN)                 - o *
                          | Título propiamente dicho               - o *
                          | Título paralelo                            *
                          | Tíitulos alternativos                       *
                          | Letras de clasificación                  o *
                          | Números de clasificación                 o *
                          | Cutter de autor                          o *
                          | Cutter de título                         o *
                          | Año de constitución                      o *
                          | Otra información sobre el título           *
                          | Area cronológica                           *
    Titulos en acervo     | Mención de responsabilidad de la edición   *
                            Menciones de responsabilidad subsiguientes*
                          | relacionadas con edición
                          | Edición                                  o *
                          | Lugar de publicación                       *
                          | Editorial                                o *
                          | Año de publicación                         *
                          | Descripcion física                         *
                          | Material ilustrativo                       *
                          | Serie                                      *
                          | Notas                                      *
                          | ISBN                                       *
                          | ISSN                                       *
                          | Clave-autor                            -   *
                          | Clave-tema                                 *
                          | Clave empleado                             *
                          | Precio unitario                        - o
                          | Fecha de compra                        - o
```

Los atributos marcados con ~ seran utilizados por la sección de Selección y Adquisiciones, los marcados con o por la sección de Biblioteca y los marcados con * por la sección de Análisis Bibliografico.

El mismo procedimiento se hace para cada entidad de las Bases de Datos lógicas de los diferentes modulos que integran al sistema. Existiran casos en los que la información de una entidad no sea utilizada por la totalidad de los modulos.

Se pueden implementar los modulos por separado, pero los archivos que se utilicen deberán contener en la definición de la estructura de la información, la que se derivó para el Sistema de Información Integrado. De esta forma se tiene un diseño integral en la conceptualización e independencia y modularidad en el desarrollo e implantación del Sistema.

Una vez que se tiene completo el Diseño Lógico de la Base de Datos del Sistema, se procede a un análisis retrospectivo en el que se debe considerar lo siguiente:

1. Revisar la factibilidad de generar cada uno de los reportes y pantallas a traves de la información definida en las distintas entidades.
2. Si faltó algún dato, se debe agregar a la entidad correspondiente (puede ser alguna clave para relacionar información)
3. Se identifican los casos conflicto que impliquen la apertura secuencial de archivos y búsqueda de información en más de tres archivos, ocasionando que el sistema sea lento. Se evalúa la posibilidad de incluir algun atributo redundante que disminuya el tiempo de

acceso a la información en consultas o emisión de reportes. En el caso de Minisis son las uniones DS requeridas.
4. Se normaliza la Base de Datos Lógica.

Diseño físico de la Base de Datos

A partir de la Base de Datos Lógica normalizada del sistema, se procede al diseño fisico de la Base de Datos, el cual es dependiente del equipo y del Administrador de Bases de Datos a utilizar.

En este caso, usando Minisis, se evalúa que archivos serán de tipo Master RD o Ksam en función del tipo de datos y uso de la información. Se identifican las definiciones Ps y Ds necesarias teniendo en mente las consultas y reportes requeridos por el usuario.

De manera ilustrativa considérense las entidades: Títulos en Acervo, Temas, Autores y Ejemplares de la Base de Datos lógica del Sistema. (ver anexo 2) Observando el contenido de información de la entidad Títulos en acervo, se observa que el tamaño de los datos es muy variable, por lo que es recomendable el manejo de memoria dinámica. Para cada título se pueden tener diferentes autores y varios temas, por lo que serán campos repetibles. Esto sugiere declarar esta entidad como una base Master RD. Se evaluó la conveniencia de substituir el nombre de autor y el tema por las claves indicadas en la entidad títulos en acervo, pero se consideró conveniente dejarlo en forma externa como se obtuvo en la Base de Datos lógica, para de esta forma poder obtener los catálogos de autoridad.

Existen varios ejemplares de un titulo, esto sugeriria que se incluyera como campo repetible y subdividido la entidad Ejemplares en la Base Títulos en acervo. Sin embargo, el movimiento de información es muy grande en relación a los datos de un ejemplar. Lo que implicaría que se tendria que correr frecuentemente procesos con Garbage. Si se adhiere a la base títulos en acervo, el proceso de garbage sería muy lento, por lo anterior se decidio que la entidad Ejemplares fuese una base independiente del tipo Master Rd. Y se def inió la union DS de Ejemplares y Títulos en acervo mediante el número de adquisición y clave del titulo, como lo sugiere la B.D. lógica.

En resúmen, es conveniente definir

Base RD Master, a la entidad que contenga información de longitud muy variable, campos repetibles y subdivididos, preferentemente que se realicen pocas modificaciones en el contenido de la información.

Ksam, archivo con registros que contienen datos de tamaño poco variable, dinámicos. No deben de existir campos repetibles y preferentemente no subdivididos. Frecuentemente es usado para validar información.

PS, definición de proyección de información, son función de la vista de información que requiere el usuario. El contenido de los archivos PS se obtiene, fácilmente considerando la información que se requiere a través de las entidades de las Bases de Datos de cada módulo y de los reportes y consultas definidos por el usuario.

DS, definición de uniones de archivos. Esta definición es necesaria para relacionar información de dos o más Bases de Datos. Se obtienen al llevar el diseño de la Base de Datos lógica al diseño de la Base de Datos física y de evaluar las uniones que se requieren para satisfacer las consultas o reportes solicitados por el usuario.

Una vez definidos los tipos de archivo y bases que se tendrán hay que completar el diccionario lógico para definir cada campo de acuerdo a las especificaciones solicitadas por Minisis, como son si el campo es repetible, indexado, obligatorio, etc.

Con el diseño físico de la Base de Datos se realiza el análisis de eventos, para determinar como se da de alta, baja o modifica el contenido de los registros de los archivos de la base de datos y se evalúa que procesos se llevaran directamente por Minisis y cuales sera necesario desarrollar mediante programación externa.

Conclusiones.

La participación del usuario (personal de la biblioteca) es constante desde el análisis. El usuario al pensar en el tipo de reportes que requiere, el ordenamiento y selección de la información, reconsidera procedimientos y no carga al sistema con información que no sera utilizada. El usuario sabe que esperar del sistema, que opciones de consulta y generación de reportes tendra.

Disminuye el rechaso a la automatización al sentirse participe del diseño y desarrollo del sistema, el cual le facilitara y ayudara en su labor cotidiana.

La metodología de Warnier basada en las salidas, permite que en forma sistemática se contemple toda la información que debe de estar en la Base de Datos. Esta metodología ayuda a evitar redundancias de información inecesarias y permite un diseño integral, pero a su vez los modulos son independientes y se puede desarrollar el Sistema por etapas.

Se puede obtener de manera sencilla e inmediata el diseño de la Base de Datos física en Minisis, a partir de la Base de Datos Lógica obtenida en la Metodología de Warnier.

El sistema diseñado podra ser utilizado por otras bibliotecas de Instituciones educativas con características similares.

En función de las salidas del sistema, el usuario de otras bibliotecas puede evaluar si se adapta a sus necesidades algun modulo o la totalidad de estos, o bien usar la definición de las salidas de base para el rediseño de reportes y consultas.

El personal de cómputo podra evaluar en forma sencilla si los reportes nuevos o las modificaciones se pueden generar directamente desde Minisis, usando la estructura de la Base de Datos ya existente o si es necesario hacer cambios en la definición de la estructura de la información.

Si desea profundizar en la metodología expuesta se recomienda consultar la siguiente bibliografía:

Logical constructions of systems/c 1981 Ed. New York, Van Nostrand Reinhold Warnier, Jean Dominique

Logical construction of programs/c 1974 Ed. New York Van Nostrand Reinhold 3 ed. Warnier, Jean Dominique

En particular sobre diseño de consultas de catálogos en linea.

Bibliographic Displays in the Online Catalog By Walt Crawford Ed. Knowledge Industry Publications, lnc. (KIPI) White Plains, NY and London

SECCION : BIBLIOTECA

CLAVE : RBI001.

TITULO : Generación de papeletas de reclamo.

OBJETIVO : Generar las papeletas de reclamo para usuarios con adeudo de material, y controlar la devolución de los mismos.

MEDIO DE SALIDA : Impresora.

ORDENAMIENTO : División, matricula.

PERIODICIDAD : Trimestral (12a. sem. del trimestre).

FORMA : Papel stock de 11" × 15".

SELECCION : Alumno, empleado.

VOLUMEN ESTIMADO : 200 páginas.

DISTRIBUCION

USUARIO	1
Coordi. Sistemas Escolares	2

SECCION : BIBLIOTECA

CLAVE : RBI001.

CONTENIDO :

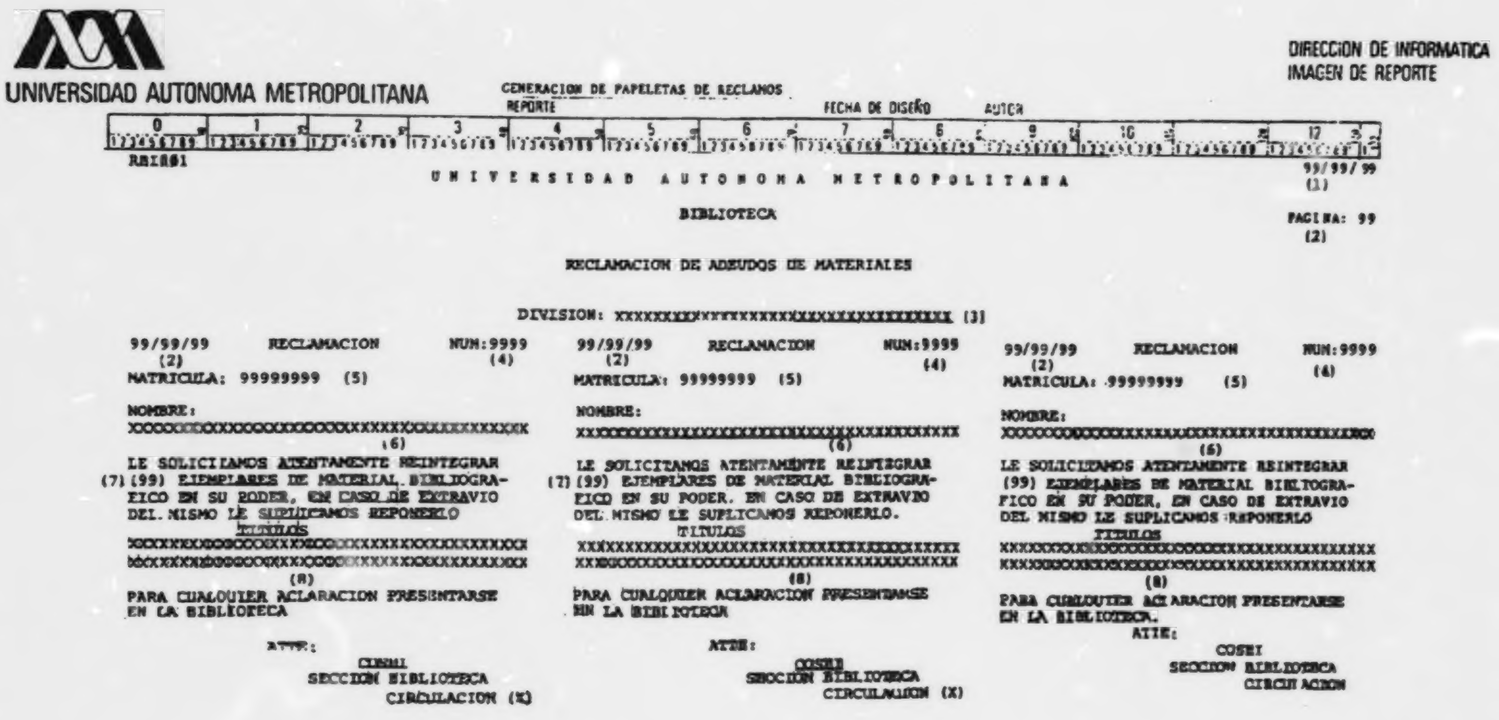
SECUENCIA	DATO	FORMATO	ORIGEN	EXPLICACION
1	FECHA	AA/MM/DD	D	Es la fecha en la que se imprime la ficha de reclamación. Se toma del sistema.
2	PAGINA	9(2)	D	Indica el número de página que se imprime para el reporte. Se deriva de un contador que se incrementa cada que hay un cambio de página.
3	DIVISION	X(35)	F	Es la organización académica de las unidades universitarias.
4	NUMERO DE FICHA	9(4)	D	Indica el número de ficha de reclamación que se imprime fisicamente. Es un número secuencial y se deriva de un contador de fichas de reclamo.
5	NOMBRE	X(40)	F	Es el nombre de un usuario alumno de la biblioteca para el que se genera la ficha de reclamación.
6	MATRICULA	9(8)	F	Es el número que se le asigna a un alumno y que lo identifica de manera única.

SECCION : BIBLIOTECA

CLAVE : RBI001.

CONTENIDO :

SECUENCIA	DATO	FORMATO	ORIGEN	EXPLICACION
7	NUMERO DE EJEMPLARES QUE ADEUDA.	9(2)	F	Indica el número de ejemplares que el alumno adeuda a la biblioteca.



UNIVERSIDAD AUTONOMA METROPOLITANA

DIRECCION DE INFORMATICA
IMAGEN DE REPORTE

GENERACION DE PAPELETAS DE RECLAMOS
REPORTE

FECHA DE DISEÑO

AUTOR

```
RBIR01
                  U N I V E R S I D A D   A U T O N O M A   M E T R O P O L I T A N A                                99/99/99
                                                                                                                      (1)
                                               BIBLIOTECA
                                                                                                                  PAGINA: 99
                                                                                                                      (2)
                                    RECLAMACION DE ADEUDOS DE MATERIALES

                                      DIVISION: XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX (3)

99/99/99      RECLAMACION       NUM:9999  99/99/99      RECLAMACION       NUM:9999  99/99/99      RECLAMACION       NUM:9999
  (2)                             (4)       (2)                             (4)       (2)                             (4)
MATRICULA: 99999999  (5)                  MATRICULA: 99999999  (5)                  MATRICULA: 99999999    (5)

NOMBRE:                                   NOMBRE:                                   NOMBRE:
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
                          (6)                                       (6)                                       (6)
LE SOLICIIAMOS ATENTAMENTE REINTEGRAR     LE SOLICITAMOS ATENTAMENTE REINTEGRAR     LE SOLICITAMOS ATENTAMENTE REINTEGRAR
(7) (99) EJEMPLARES DE MATERIAL BIBLIOGRA-  (7) (99) EJEMPLARES DE MATERIAL BIBLIOGRA-  (99) EJEMPLARES DE MATERIAL BIBLIOGRA-
FICO EN SU PODER, EN CASO DE EXTRAVIO     FICO EN SU PODER. EN CASO DE EXTRAVIO     FICO EN SU PODER, EN CASO DE EXTRAVIO
DEL MISMO LE SUPLICAMOS REPONERLO         DEL MISMO LE SUPLICAMOS REPONERLO.        DEL MISMO LE SUPLICAMOS REPONERLO
          TITULOS                                   TITULOS                                   TITULOS
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
                (8)                                       (8)                                  (8)
PARA CUALOUIER ACLARACION PRESENTARSE     PARA CUALQUIER ACLARACION PRESENTARSE     PARA CUALQUIER ACLARACION PRESENTARSE
EN LA BIBLIOTECA                          EN LA BIBLIOTECA                          EN LA BIBLIOTECA.
                                                                                                  ATTE:
          ATTE:                                     ATTE:                                              COSEI
                 COSEI                                     COSEI                                   SECCION BIBLIOTECA
            SECCION BIBLIOTECA                        SECCION BIBLIOTECA                                   CIRCULACION
                    CIRCULACION (X)                           CIRCULACION (X)
```

ENTIDADES DE LA BASE DE DATOS LOGICA

Titulos en acervo del Sistema	Clave del titulo (ISN)	-	o	*
	Titulo propiamente dicho	-	o	*
	Titulo paralelo			*
	Titulos alternativos			*
	Letras de clasificacion		o	*
	Numeros de clasificacion		o	*
	Cutter de autor		o	*
	Cutter de titulo		o	*
	Ano de constitucion		o	*
	Otra informacion sobre el titulo			*
	Area cronologica			*
	Mencion de responsabilidad de la edicion			*
	Menciones de responsabilidad subsiguientes relacionadas con edicion			*
	Edicion		o	*
	Lugar de publicacion			*
	Editorial		o	*
	Ano de publicacion			*
	Descripcion fisica			*
	Material ilustrativo			*
	Serie			*
	Notas			*
	ISBN			*
	ISSN			*
	Clave-autor	-		*
	Clave-tema			*
	Clave empleado			*
	Precio unitario	-	o	
	Fecha de compra	-	o	
Temas	Clave del tema			
	Nombre del tema			
Autores	Clave del autor			
	Nombre del autor			
Trata temas	Clave del tema			
	Clave del titulo			
Tiene autores	Clave del autor			
	Clave del titulo			

CONSIDERACIONES PARA LA BASE de DATOS FISICA DEL SISTEMA

El titulo propiamente dicho, clave de autor y clave de tema seran campos indexados, que se usaran como llaves alternas de busqueda directa de informacion.

Clave de autor y clave de tema seran campos repetibles

B.D. master tipo RD

Las entidades Temas y Autores se definiran como Master RD para obtener los catalogos de autoridad.

La B.D. logica sugiere la definicion " DS " para poder relacionar Titulos en acervo, Temas y Autores.

Trata temas = Temas + Titulos en acervo
PS Trata temas de:

- Clave titulo
- Titulo propiamente dicho
- Editorial
- Edicion
- Ano de publicacion
- Numero de clasificacion
- Cutter de autor
- Cutter de titulo
- Ano de constitucion
- Precio unitario
- Fecha de compra

ENTIDADES DE LA BASE DE DATOS LOGICA

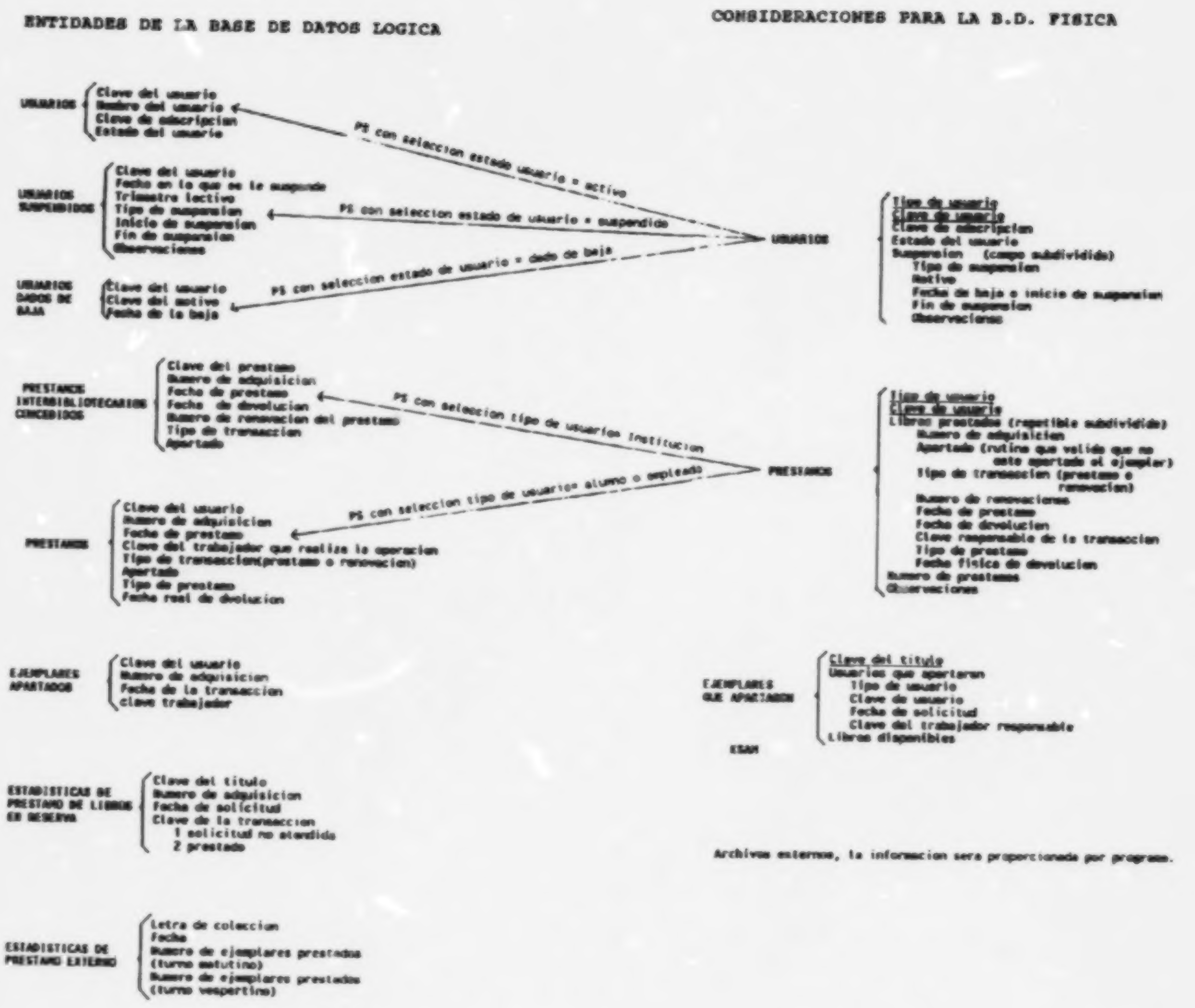
- Ejemplares que cambian de coleccion
    - Numero de adquisicion
    - Coleccion a la que pertenecia
    - Fecha en la que cambio
- Ejemplares extraviados sujetos a reposicion
    - Clave del usuario
    - Numero de adquisicion
    - Fecha de vencimiento de la reposicion
    - Tipo de reposicion
    - Observaciones
    - Fecha real de la reposicion
- Ejemplares
    - Numero de adquisicion
    - Clave del titulo
    - Estado del ejemplar
    - Tipo de coleccion
- Ejemplares con problenas
    - Numero de adquisicion
    - Tipo de problema
    - Fecha de inicio
    - Fecha de fin
    - Clave-trabajador
- Libros en encuadernacion
    - Numero de adquisicion
    - Clave del usuario
    - Clave del responsable
    - Fecha de inicio de encuadernacion
    - Fecha de fin de encuadernacion

CONSIDERACIONES PARA LA BASE de DATOS FISICA DEL SISTEMA

Dado que es posible tener campos subdivididos y repetibles, así como manejo de memoria dinamica en Minisis. Es conveniente tener un archivo Master RD de ejemplares.

- Ejemplares
    - Numero de adquisicion
    - Clave del titulo
    - Tipo de coleccion
    - Edo del ejemplar
    - Descripcion del estado (subdividido)
        - Fecha del reporte
        - Fecha de la solucion del problema
        - Clave del usuario responsable de la reposicion
        - Tipo de usuario
        - Fecha de vencimiento de la reposicion
        - Clave del trabajador
    - Cambios de coleccion (Repetible subdividido)
        - Coleccion a la que pertenecia
        - Fecha en la que cambio de coleccion

Las entidades Ejemplares que cambiande coleccion, Ejemplares extraviados sujetos a reposicion, Ejemplares con problemas y Libros en encuadernacion; Seran proyecciones "PS" de la base ejemplares descrita, los cuales tendran como campos los atributos indicados en cada entidad.


ENTIDADES DE LA BASE DE DATOS LOGICA
CONSIDERACIONES PARA LA B.D. FISICA
USUARIOS
Clave del usuario
Nombre del usuario
Clave de adscripcion
Estado del usuario
PS con seleccion estado usuario = activo
USUARIOS SUSPENDIDOS
Clave del usuario
Fecha en la que se le suspende
Trimestre lectivo
Tipo de suspension
Inicio de suspension
Fin de suspension
Observaciones
PS con seleccion estado de usuario = suspendido
USUARIOS DADOS DE BAJA
Clave del usuario
Clave del motivo
Fecha de la baja
PS con seleccion estado de usuario = dado de baja
USUARIOS
Tipo de usuario
Clave de usuario
Clave de adscripcion
Estado del usuario
Suspension (campo subdividido)
Tipo de suspension
Motivo
Fecha de baja o inicio de suspension
Fin de suspension
Observaciones
PRESTAMOS INTERBIBLIOTECARIOS CONCEDIDOS
Clave del prestamo
Numero de adquisicion
Fecha de prestamo
Fecha de devolucion
Numero de renovacion del prestamo
Tipo de transaccion
Apartado
PS con seleccion tipo de usuarios institucion
PRESTAMOS
Clave del usuario
Numero de adquisicion
Fecha de prestamo
Clave del trabajador que realiza la operacion
Tipo de transaccion(prestamo o renovacion)
Apartado
Tipo de prestamo
Fecha real de devolucion
PS con seleccion tipo de usuarios alumno o empleado
PRESTAMOS
Tipo de usuario
Clave de usuario
Libros prestados (repetible subdividido)
Numero de adquisicion
Apartado (rutina que valida que no este apartado el ejemplar)
Tipo de transaccion (prestamo o renovacion)
Numero de renovaciones
Fecha de prestamo
Fecha de devolucion
Clave responsable de la transaccion
Tipo de prestamo
Fecha fisica de devolucion
Numero de prestamos
Observaciones
EJEMPLARES APARTADOS
Clave del usuario
Numero de adquisicion
Fecha de la transaccion
clave trabajador
EJEMPLARES QUE APARTARON
Clave del titulo
Usuarios que apartaron
Tipo de usuario
Clave de usuario
Fecha de solicitud
Clave del trabajador responsable
Libros disponibles
ESAM
ESTADISTICAS DE PRESTAMO DE LIBROS EN RESERVA
Clave del titulo
Numero de adquisicion
Fecha de solicitud
Clave de la transaccion
1 solicitud no atendida
2 prestado
ESTADISTICAS DE PRESTAMO EXTERNO
Letra de coleccion
Fecha
Numero de ejemplares prestados
(turno matutino)
Numero de ejemplares prestados
(turno vespertino)
Archivos externos, la informacion sera proporcionada por programa.

MINISIS : An Experience in Health Sciences Library CAHSL Subsystem

TANG ZHAOBI
YUAN ZUPING

(CAHSL – Cataloguing and Acquisition Subsystem for the Health Sciences Library of Chinese Academy of Medical Sciences)

1. Introduction of CAMS Library

The library of the Chinese Academy of Medical Science (CAMS) was established 70 years ago, and is one of the health sciences libraries in China with a long history and a rich collection.

The library is a professional health sciences library with a collection of more than 450,000 books and journals. The library is also the medical literature center of the CAMS and PUMC, making it an important component for medical research and teaching. The functions of the library are collecting, preserving, and processing medical literature concerning the prevention and treatment of disease and research work in the medical sciences, running the eight-year system medical university, and providing service for readers.

2. Library Automation

CAMS library is using MINISIS for its various library functions, such as acquisitions, cataloguing and searching (on-line retrieval.) Other non-library application, such as the top management personnel information system will be created in the future. The CAHSL subsystem is a part of the integrated library management system. Owing to the limited hard disc space, other subsystems, such as the periodical and circulation subsystems, have not been created. This paper details the CAHSL subsystem.

3. System Configuration

At Present, five terminals have been installed at the library. These terminals are linked with the central computer room. The present system configuration is as follows:

Hardware configuration

Processor: HP3000 Series 37/XE with 2 megabytes of main memory.
Disk driver : 1*7914 = 132 megabytes
Disk driver : 2*7937 = 1148 megabytes (uninstalled)
Tape Driver : 1*7974 (1600 BPI)
Console : 2392A Terminal

4. The CAMS library automation group

An automation group was set up in the CAMS Library to implement and execute various CAMS Library automation activities. The major tasks of the group are to coordinate, design various library computerized information systems, maintain system development, conduct or coordinate training for library staff and end users on the use of CAMS library's data bases.

5. MINISIS DATA BASE in CAMS Library

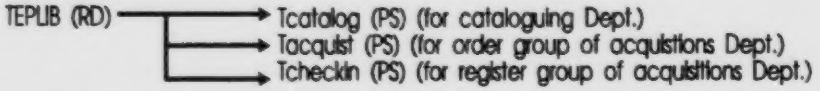
At present, there are four MINISIS data bases in CAMS library of which are listed in the following table:

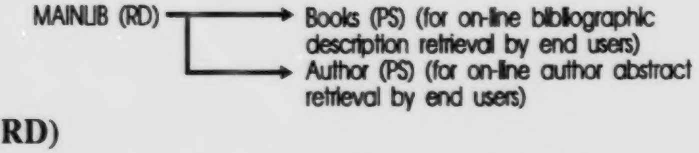
NAME	USE
MAINLIB	STORE THE LIBRARY'S BOOK COLLECTION AND BE USED AS ON-LINE BIBLIOGRAPHIC RETRIEVAL DATA BASE FOR END USERS.
TEPLIB	STORE THE CATALOGUING AND ACQUISITION INFORMATION TO BE USED TO PRODUCE STANDARD BIBLIOGRAPHIC RECORDS AND VARIOUS STATISTICS.
LM88	STORE THE PERIODICAL UNION CATALOGUE FROM 1985-1990.
CNMEDLAR	CHINESE BIOMEDICAL LITERATURE ANALYSIS AND RETRIEVAL SYSTEM (IN TESTING PHASE)

6. CAHSL subsystem structure and function

CAHSL consists of two RD data bases and five PS data bases. The relationships between them are shown in following illustration.

Details on the functions of these data bases are elaborated in the following sections of the paper:


TEPLIB (RD)
Tcatalog (PS) (for cataloguing Dept.)
Tacquist (PS) (for order group of acquistions Dept.)
Tcheckin (PS) (for register group of acquisitions Dept.)


MAINLIB (RD)
Books (PS) (for on-line bibliographic description retrieval by end users)
Author (PS) (for on-line author abstract retrieval by end users)

Teplib (RD)

This data base consists of three parts of information, which are collected from the ordering, registration and cataloguing processes. This RD has two main purposes. One is to produce the Chinese standard bibliographic description records of new books, which are not available for searching until they are transferred to Mainlib (RD). The other is to store acquisitions and cataloguing information, as well as produce various statistics which the library routine needs.

Three projected subsets are generated from it. Each of them are derived from Teplib's data, and can share data with each other. They are the Tacqsit projected subset for ordering, the Tcheckin projected subset for registration, and the Tcatalog projected subset for cataloguing.

Tacqsit: A projected subset data base with write access capability which is used to input book ordering data, such as the order date and the estimated price, to get the primary data of the ordered book from the order list, and to produce the financial statistics of orders.

Tcheckin: A projected subset data base with write access capability which is used to input the data of registration and confirmation of the price of the new book. According to the new book information, wrong information entered during the ordering process is rectified and more information on the new book which the registration process needs is added. Data is entered and updated according to the "Chinese Descriptive Cataloguing Rules for Western Language Materials." The records formed are also the part of standard bibliographic description. So, part of cataloguing work is finished in this data base. Tcheckin is also used to produce the accurate cost statistics as well as other statistics which the Acquisitions Dept. needs.

Tcatalog: A projected subset database with write access capability which consists of the bibliographic data shared with Tcheckin and other cataloguing information. It is used to complete cataloguing work, such as adding subject headings, classification numbers and some collection management information. A completed record of standard bibliographic descriptions is formed and a standard bibliographic card is produced from this data base.

Mainlib(RD)

This data base stores the Chinese standard bibliographic description of the library collection available for loan. The records of this RD are converted from Teplib RD through ISO MODEL, which can not be updated by except by an assigned high level Librarian. With ISO 2709 MARK TAPE, it can exchange data with other data bases. Mainlib is used as the on-line bibliographic description retrieval system, for readers and library staff. To provide a more convenient on-line service for the end users, two projected subsets are generated from it.

Books (PS): A projected subset without write access capability where end users are only allowed to query information on books already catalogued and available for loan in the library.

Author(PS): A projected subset without write access capability where end users are allowed to query the information about the author.

ILL. 1 shows the structure and function in detail.

7. Characteristics of CAHSL

1. Definition of Teplib and Mainlib RD data bases:

The Health Sciences Library of Chinese Academy of Medical Sciences is the head Health library in China. So, the quality of cataloguing required is higher and more standardized than other libraries. How could the field definition meet the requirement of standard cataloguing rules, and be suitable for the CAMS library? After comparing the UNIMARK definition, the LCMARK definition, and the CATLINE data base definition, we choose to define the field definitions of the two RD data base, Mainlib and Teplib, following the model of CATLINE. It is not only very suitable for the library, it also brings a big advantage — a standard data resource.

CATLINE (CATalog-onLINE) contains approximately 600,000 bibliographic records related to the biomedical sciences; this is virtually all of the catalogued titles in the National Library of Medicine's collection. The subject headings used by CATLINE are from MeSH, which is used widely in health sciences libraries and is very important for health sciences libraries. Considering these characteristics, we hope to take advantage of CATLINE as a standard data resource in the data retrospection.

2. Full Screen Menu

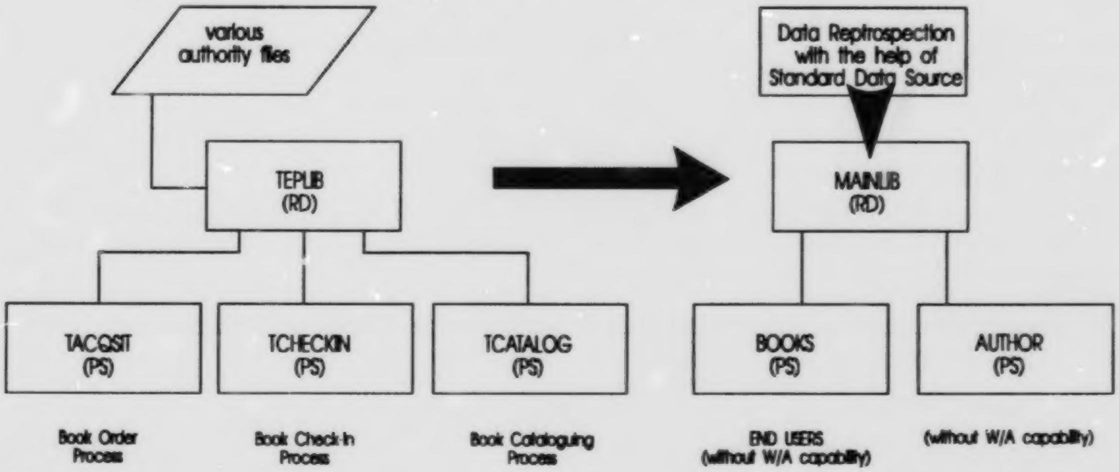
Using VPLUS, we designed a two-level full screen menu for each PS data base with the w/a(write/access) capability, which helps the staff input or update data easily.

3. Program for the batching produced bibliographic cards

In CAMS library, the catalog card has been kept for 70 years. The computer is used in the cataloguing and the catalog card is still kept complete. Accordingly, a COBOL program is used to produce bibliographic cards in batch and a Batchin file for MINISIS at the same time. This program provides a batch input/output chance.

8. Future development

The other parts of integrated management system are still in the design stage. It is hoped that should all these phases of CAMS library automation programme be completed, it will enable CAMS library to increase the efficiency of its existing data bases so that library users will benefit and fully utilise the library's resources.


various
authority files
TEPLIB
(RD)
TACQSIT
(PS)
TCHECKIN
(PS)
TCATALOG
(PS)
Book Order
Process
Book Check-In
Process
Book Cataloguing
Process
Data Reptrospection
with the help of
Standard Data Source
MAINLIB
(RD)
BOOKS
(PS)
AUTHOR
(PS)
END USERS
(without W/A capability)
(without W/A capability)

MINISIS ：在卫生科学图书馆中的应用

—— CAHSL 子系统

(CAHSL —— 中国医学科学院(CAMS)图书馆集成管理系统采编子系统)

鹰赵必　　袁祖平

一、中国医学科学院(CAMS)图书馆简介

中国医学科学院图书馆成立迄今已有七十年了，是我国历史悠久，藏书丰富的医学图书馆之一。中国医学科学院图书馆同时也是中国协和医科大学图书馆。

医科院图书馆是一所医学专业图书馆，目前馆藏书刊超过 45 万册，是中国医学科学院和中国协和医科大学的医学文献中心，是医、教、研工作的重要组成部分。其具体任务是：密切配合防病治病的需要，围绕发展医学科学的研究，为增进人民健康和办好八年制医大收集、整理国内外生物医学的文献资料，并提供各项服务。

二、 图书馆自动化

目前，我馆已经开始将 MINISIS用于图书馆的日常工作管理，如图书的采访、编目、检索。其他的非图书管理如人事管理系统等也将建立。由于受磁盘空间的限制，其他的子系统如期刊管理子系统、流通子系统尚未建立。本文将对 CAHSL子系统作详细的介绍。

三、系统配制

目前，我们有五台Vectra作为终端联接在主机上。系统的主要配制如下：

主机	HP3000 37/XE　内存：2MB
磁盘机	1 X 7914 = 132 MB
磁盘机	2 X 7937 = 1148 MB (11月份安装)
控制台	2392A

四、 CAMS图书馆自动化小组

为实现图书馆自动化管理，我馆成立了自动化小组。它的主要任务是协调和设计图书馆自动化情报管理系统，以及系统的维护和发展，指导和培训图书馆工作人员和用户对我馆数据库的使用。

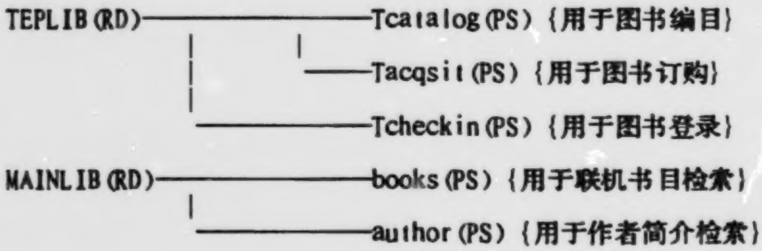
五、 MINISIS数据库

目前，我们已建立了四个 MINISIS 数据库，如下表所列：

数据库名	用途
MAINLIB	存储图书馆藏，并为用户提供联机书目检索
TEPLIB	存储图书的采编信息，用于产生标准书目及其他统计数据
LM88	存储1985-1990年全国高等医药院校图书馆期刊联合目录
CNMEDLAR	中国生物医学文献数据库（在设计中）

六、 CAHLS 子系统的结构和功能

CAHLS 包括两个RD数据库和五个PS数据库。它们之间的关系如下图所示：


TEPLIB(RD)
Tcatalog(PS)｛用于图书编目｝
Tacqsit(PS)｛用于图书订购｝
Tcheckin(PS)｛用于图书登录｝
MAINLIB(RD)
books(PS)｛用于联机书目检索｝
author(PS)｛用于作者简介检索｝

下面就上述各数据库的功能作详细介绍：

TEPLIB(RD)：

这个数据库包括三个方面的信息，分别来自订购、登录和编目处理过程。它的主要任务有两个：其一是用来产生新书具有中国标准书目著录格式的书目记录（这些记录只有传送到MAINLIB(RD)库中后，才供读者借阅）；另一个任务是存储采编信息，以产生图书馆业务所需的各种书目卡片和统计资料。

在这个RD库下，有三个投影库。每个子库均为主库TEPLIB的一部分投影，通过主库它们相互可共享一些数据，但又有各自独用的数据。这三个投影库分别是：供图书订购用的TACQSIT；供图书登录用的 TCHECKIN 和供图书编目用的 TCATALOG。它们都是具有写能力的投影数据库。

TACQSIT：这个PS库用于输入图书订购的数据，如：订购日期、图书估计，及根据征订目录输入所订图书的基本信息。同时亦产生图书订购单和订购书价统计表。

TCHECKIN：这个 PS 库用于输入登录数据，如：确定新书的价格，根据新书修改订购时输入的数据，增加一些登录所需的新内容。由于书目是按照《中国西文文献著录条例》输入的，因此输入的数据具有标准的著录格式，也就是说部分的编目工作在此已经完成。它同时也产生精确的书价统计及其他采访部门需要的统计数据。

TCATALOG：这个PS库包含了TCHECKIN里的文献著录数据，它用于完成文献的编目加工，诸如主题标引、分类及输入一些馆藏信息。这样，完整的标准文献著录工作就完成了，利用这个数据库，可产生各种标准书目卡片和馆藏统计资料。

MAINLIB(RD)：

这个数据库用于存储图书馆所有可供借阅的图书文献书目。它的书目记录是从TEPLIB数据库传送来的。除了指定的高级馆员外，这个数据库对其他终端用户是只读的。由于它的数据库结构与CATLINE数据库的结构相似，因此可以与CATLINE数据库兼容。它还可以利用ISO-2709格式磁带与其他数据库进行数据交换。它为图书馆工作人员和读者提供联机书目检索系统。为了提供更有效的服务，我们从这个数据库中引出了两个投影库，它们都是只读PS库。

BOOKS(PS)：它为读者和工作人员提供可供借阅的图书文献的联机书目检索。

AUTHOR(PS)：它为读者和工作人员提供作者简介的检索。

图一展示了各数据库的结构、功能和相互关系。

七、 CAHLS的特点

1. TEPLIB 和 MAINLIB 数据库的定义

鉴于中国医学科学院卫生科学图书馆居中国卫生科学图书馆之首，所以，图书编目的质量比其他馆要求更高、更标准化。如何使数据库字段的定义既符合标准著录的规则，又适合我们图书馆？在对UNIMARK、LCMARK、CATLINE等数据库定义的研究和比较之后，我们

对两个RD数据库TEPLIB和MAINLIB的字段定义选择了与CATLINE相似的定义。它不仅非常适合于我们图书馆，同时在各方面给我们带来了极大的便利 — 标准的数据源和广泛的兼容性。

CATLINE(CATalog-onLINE)贮存有将近600 000条生物医学方面的书目记录，它实际上包括了美国国家图书馆所有已编目的医学文献。它使用的主题词表是在卫生科学图书馆中使用最广泛、也是最重要的MASH主题词表。基于这些特点，我们希望能使用 CATLING作为数据回溯的标准数据源。

2．全屏幕编辑

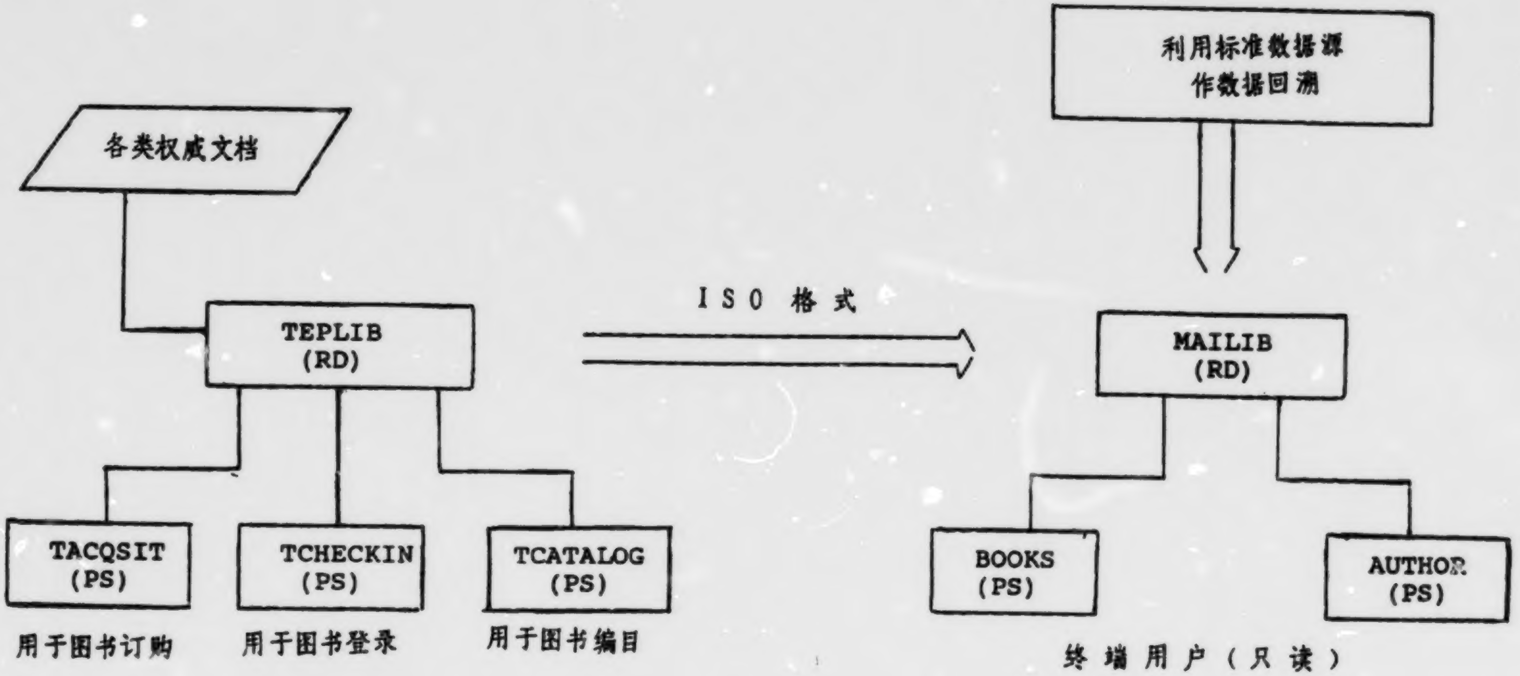
对每个具有写能力的PS数据库，我们都利用VPLUS设置了二级全屏菜单。这样，工作人员能很容易地输入或修改数据。

3．使用程序产生书目卡片

在CAMS图书馆，编目卡片已有70年的历史了。虽然开始使用计算机进行图书管理，但是，图书馆还是希望保持书目卡片的完整性。因此，我们编制了一个 COBOL 程序，用于成批生产书目卡片，同时此程序也产生由 TEPLIB 数据库向 MAINLIB 数据库传送新书数据的 MINISIS 的 BATCHIN 文件。

八、 展望

图书馆集成管理系统的其他部分还在设计之中，希望在不久能将整个管理系统完善，并正式投入应用，从而充分利用图书馆现有的资源，以发挥更大的社会效益。


各类权威文档
TEPLIB (RD)
TACQSIT (PS)
TCHECKIN (PS)
TCATALOG (PS)
用于图书订购
用于图书登录
用于图书编目
ISO 格式
利用标准数据源
作数据回溯
MAILIB (RD)
BOOKS (PS)
AUTHOR (PS)
终端用户（只读）

图一、 CAHSL 子系统结构图

INTEGRATED SYSTEM OF PAO SUI-LOONG LIBRARY

ZHEN QIAO-YING
YAN CHUN
Pao Sui-loong Library of Shanghai Jiao Tong University

KEYWORDS

COMPUTER, LIBRARY, IMAGE, MINISIS, INTEGRATED SYSTEM

ABSTRACT

MILIS is a practical multi-user library integrated system which was developed by Pao Sui-Loong library. It can be run in HP3000 series computer. MILIS based on MINISIS database management system supplied by IDRC, and IMAGE DBS attached to HP3000 computer. It consists of following six subsystems: 1) Acquisition; 2) Cataloguing; 3) Serials; 4) Circulation; 5) Financial; 6) Public access

I. INTRODUCTION

MILIS is a practical multi-user integrated online system which was developed by Pao Sui-loong library. MILIS is an acronym for MINISIS and IMAGE Library Integrated System. It can be used in HP3000 series computer. MILIS based on MINISIS database management system supplied by IDRC (International Development Research Center), and IMAGE database management software attached to HP3000 computer. A lots advanced technology products in computer field such as CD-ROM, light pen, bar code and simulation technology are also adopted, so the whole system can rapidly process mass jobs in the library. MILIS with standard bibliographic records can include the all fundamental works in the middle or large scale libraries. This will be a base of Pao Sui-loong Library computerized management.

II. SYSTEM ANALYSIS

The main idea of MILIS is set to a library integrated management system which contains daily work in library such as acquisition, cataloguing, circulating, serials controlling, financial controlling and public accessing, etc. The main features in MILIS are integration, practicality, standardization of bibliographic records and resource-sharing.

Integrity in MILIS means to face all collection in western language and Chinese, makes acquisition cataloguing, circulation and financial control as one, and enhances the bibliographic data transfer ability among these subsystems. Every subsystem has own procedure data. The data related to several procedures will be put into sharing data set. For example, some bibliographic data used in circulation subsystem can be transferred from projected subset of main bibliographic database. Using the same method, we can transfer financial managing data from subsystems of ordering (including books and periodicals).

Implication of practicality is to form a friendly interface between MILIS and users such as full screen input pattern and online help. The system is easy to use, secure, reliable. There are four-step security measurement and system back-up. Due to simulation technology in workstations some jobs in MILIS can work normally such as circulation subsystem, if the main frame is down. It can be connected with former manual system in libraries. MILIS can embody the advantages of automatic management system which could not be done in traditional method.

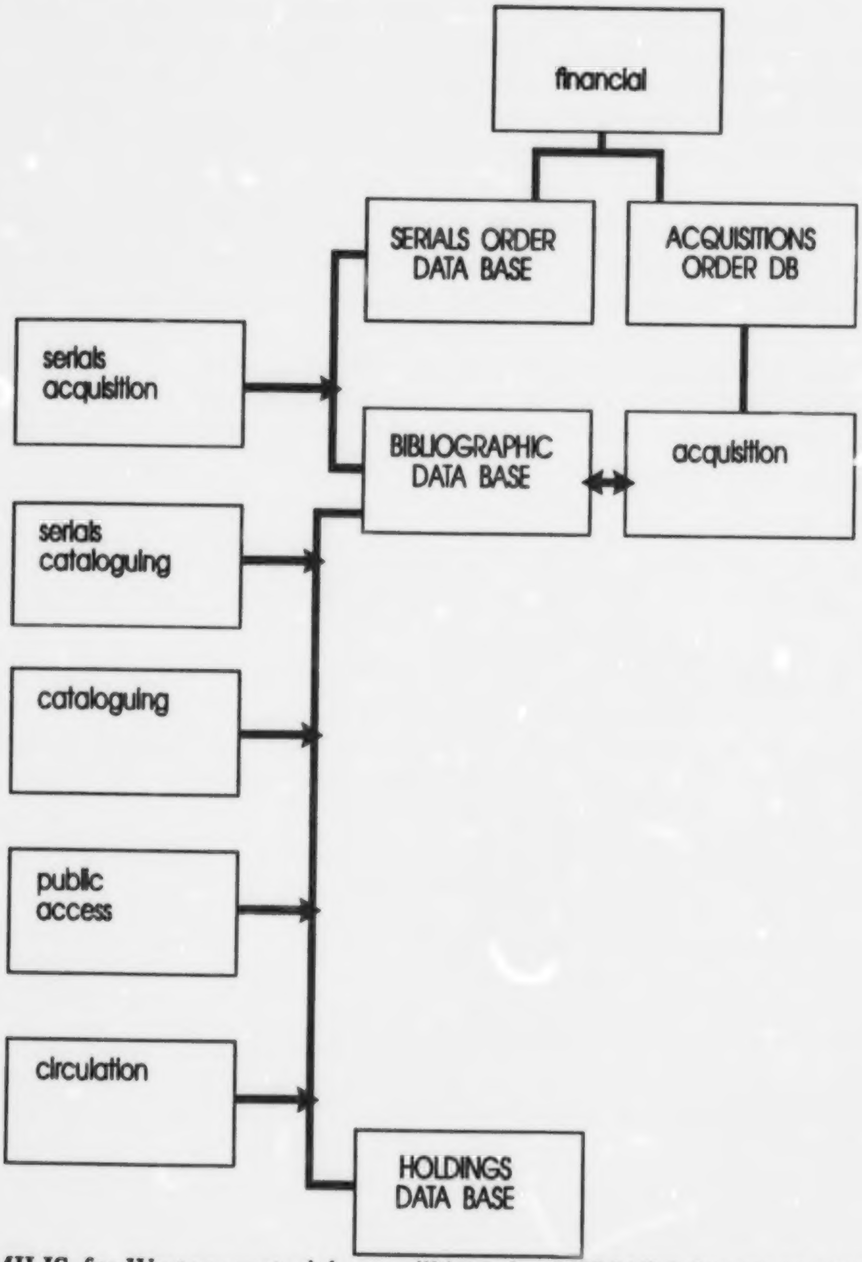
The bibliographic data needed in this system must be standardization. For western materials, we catalogue with AACR2, for Chinese with CATALOGUING RULES OF CHINESE DOCUMENT Data storing formats in MILIS are USMARC and UNIMARC.

The one of the advantages using MILIS integrated system is resource-sharing. The standard USMARC records from the Library of Congress and standard Chinese records from Chinese National Library can be directly accepted to MILIS through the CD-ROM disc or the tapes.

After analyzing the characteristics of MINISIS's and IMAGE's DMS, we built sub-systems of acquisition, cataloguing, serials control and public access in MINISIS, because the relative DB of MINISIS's DMS, which can support varied fields, is suitable to process fundamental works of library and has strong retrieval function. IMAGE is a network DB, the DMS has a quicker response speed, so we set up circulation subsystem and financial control subsystem in it because they need higher speed. Data communication in MINISIS and IMAGE can be executed through the interface between them. For improving whole performance of this system, we allocate some new technologies to MILIS, using CD-ROM recorded USMARC format to be standard data resource of foreign materials, light pen and bar-code to raise accuracy of input information and speed of lending, link of microcomputer and mainframe to transmit media data from microcomputer to HP3000 computer, simulation technology to substitute some functions in MILIS when host computer does not work.

Full integrated system after modular design may strengthen flexibly. The subsystems can be collected as a integrated system and can independently perform some indicated functions.

III. RELATIONSHIP BETWEEN SUBSYSTEMS AND MAIN DATA BASE IN MILIS


financial
SERIALS ORDER DATA BASE
ACQUISITIONS ORDER DB
serials acquisition
BIBLIOGRAPHIC DATA BASE
acquisition
serials cataloguing
cataloguing
public access
circulation
HOLDINGS DATA BASE

Note: In MILIS, for Western materials we will be using USMARC format and for Chinese materials we will be using UNIMARC format from the Beijing Library.

VI. INTRODUCTION TO MILIS

I. Acquisition subsystem

In acquisition department, book information is always supplied by publishers and book traders. The incomplete and unreliable data in acquisition database are only procedure data in whole integrated system. In order to keep data reliable in main DBs and avoid to confuse data with other DBs, acquisition subsystem has two independent DBs – ordering database and budget controlling DB. Different PSs from two DBs can be used according to staff need.

Main functions of acquisition subsystem are:

Query module: Full information of arriving book is kept in main bibliography DB after cataloguing so data in acquisition subsystem are procedure data. For saving internal storage space book data in acquisition DB will be kept for 3 years generally, so we must duplicate check for ordering in two DBs, one is main bibliography DB, another is acquisition DB. Major access points of this function module are: title, author, ISBN, series title, proceeding title, and publisher. Official query is used at some special conditions such as purchasing situation, fund allocation, need of printing urgent notice or not, etc.

Ordering: There are two ways to input ordering data. First is manual in-put using screen inputting format on terminal. Second is computer transmission from standard data sources such as ANY-BOOK and UNIMARC to acquisition DB and then plus order information.

Checking: Staff member draws book information by retrieving and puts it into a file, then the hitfile is set up. We use batch method to add checking contents. After checking, the system can plus arriving mark on the records. At the same time, preparing for transmit some fields which finance subsystem requests has been finished.

Handling exchanging and gift books: Handling exchanging and gift books is generally as same as book checking, but these books data must be added exchanging or gift mark in DB to meet the needs of statistics and administration.

Request analysis: For using collection better and raising quality of collection higher, we set a request analysis DB. Patron could fill requirement application on terminal. It is benefit for librarian to choose best books. System of this DB will show reader's requirements, library holdings and funds using state when he chooses books.

Varied statistics: (1) Budget statistics; (2) Arrived book statistics; (3) Publisher and vendor statistics; (4) Exchange and gift book statistics; (5) Catalogue statistics; (6) Fund using statistics.

Print reports: (1) Print order form; (2) Print new-book report; (3) Print exchange and gift report; (4) Print urgent notice; (5) Print statistics report.

2. Catalog subsystem

Duplicate check: The function supplies following search points: (1) Bar code, (2) ISBN; (3) LC card number; (4) Author; (5) Title; (6) Conference title; (7) Corporation title; (8) Subject; (9) Local call number; (10) Author and title; (Il) Serials; (12) General query.

Assistant classifying: The function provides a reference for library staff to gather different titles with the same classification number, in addition, if the record taken from standard bibliographic resource, such as LC and Dewey classification numbers is also a good reference.

Original catalog: MILIS uses full MARC format to process bibliographic data. The subsystem utilizes USMARC format for western materials and UNIMARC format for Chinese. For bibliographic record data base, bibliographic data must be standard. If there is not find records in standard resource data, a cataloguer should fill worksheets and check each fields then input them.

Automation cataloguing: Modifies other standard source bibliographic data, such as bibliographic data from BIBLIOFILE system and MARC tape, as own records. After modifying and adding some fields in a record,the records are automatically transferred to local catalog data base.

Catalog maintenance: Modifies the local catalog records which are identified with cataloguing errors or something needed to be corrected.

Catalog production: (I) Printing collection; (2) Printing newsletter; (3) Printing catalog cards: a) sorted by classification b) sorted by edit c) sorted by author d) sorted by series title e) sorted by subject f) printing set of card; (4) Printing label; (5) Printing registration

Holding management: (I) Add holdings information to new catalogued titles, such as copy number; bar code; price; stock location etc. (2) Modifies holdings information which needs to be corrected or maintain bar code.

Statistics report: (I) Collection statistics; (2) Language statistics; (3) Classification statistics.

3. Circulation subsystem

The system is a multi-user and online library circulation system with

bar coding technology and light pen. It is to be run on any computer of HP3000 series, and based on IMAGE/3000 network database management system. The circulation system not only is a subsystem of the integrated system, but also can run alone. That bibliographical information in circulation data base is directly taken from catalog subsystem by the interface MINISIS-IMAGE. The circulation subsystem mainly provides the functions as following:

Check-out: The function automatically checks patron's

qualifications including library ID card expiration, borrower's registration, patron borrowing privilege, etc.

Renewal: The function of renewal is the same as charge, in addition that it can check patron's renewal privilege.

Check-in: The discharge module will indicate the overdue book and hold book at the processing time.

Holds: First, the system automatically checks patron's validation for hold. If the patron is qualified, staff will input the call number of the hold book.

Fine payment: If the book is overdue, the system will compute the fines. The patron must pay fine then can charge the books.

Query: In query module, there are two main query functions. One is public query, and the other the staff query. It provides access points: call number; barcode number of a book; barcode number of a borrower's card; patron's name; date, etc.

Statistics: It is to print monthly or annual reports.

Overnight processing: It generates various kinds of notices for overdue books, holds, fines, etc.

Patron's file maintenance: This function permits the registering, cancelling, modifying of the patron's badges on request.

4. Serials control subsystem

(I) Acquisition

Acquisition system converts the information obtained in processing materials into local holdings through a computer-based data flow. The original data may be keying the system in selection or ordering functions. The acquisition system provides the functions as following:

Serials selection: The function can improve the relationship between patrons and librarians. Periodical selection and cancellation that occurs at the same time in library collection updating are separately treated in the different functions. Serials selection is based upon acquisition policy and organized form to collect serial materials Online selecting decision will be a good reference for ordering control.

Ordering: Ordering functions are able to print order forms, add order forms, order query and renew subscription notices, etc. At the same time, financial accounting will be associated with fiscal year budget, allocated account by linking different physical files to organize them into user-viewed, system-wide record.

(2) Cataloguing

Serials cataloguing is similar to book cataloguing in many ways. There are six functions in serials cataloguing: automatic cataloguing; original cataloguing; maintenance of online cataloguing; printing cards; holding maintenance; on-line catalog query.

(3) Check-in

Check-in section divides into current issue check-in, binding, claiming, check-in query. The automation system enables to predict new issue receiving, flag missing issues, ready to bind issues and can put them into their own data base.

The bibliography information in check-in data base is directly taken from local online catalog message merged with inputting check-in message. The system generated predication depends on the local submitted check-in criteria. Predication aids claiming, bindery control. Periodicals processed in bindery, claiming control will reflect their binding, missing, claiming status in check-in charts, that makes check-in data presents a public access usage and become an excellent path for patrons to look for current issue status.

5. Financial management subsystem

The DB of financial management system is a two level network DB supported by IMAGE/3000. This sub-system can exchange data with Acquisition and Periodical subsystem' DB in the integrated system. There are 7 kinds of functions modules in financial management subsystem.

Control module: It can check user's access ability in financial subsystem and supply varied functions for different users according to their status.

Inputting module: It can input account and lending account data of books and serials automatically, but the information of administration account and administration lending account must be input by manual.

Modify module: It can modify or delete those accounts mentioned above and revise budget amount from collected category records.

Query module: There are main access points in it, including name of payee, type of document, date of payment. It also can allow to complete query for category account and miscellaneous account.

System initiation module: It is in charge of setting and inputting budget amount of every category and creating codes corresponding to those payees.

Printing module: (1) Prints detailed account of every payment; (2) Prints collected reports; (3) Prints current account.

Collecting module: It can collect all kinds of account, show collected category which has some change recently on screen, and automatically print collected reports whose amount have exceed the budget.

6. Public access

The patrons can search bibliographical and holding information in the collection through all kinds of access points. The information may be accessed online via the following keys: Barcode; ISBN; LC call number; Author; Title; Name of conference; Name of corporation; Subject; Local call number; Author and Title; Serials; ISSN; Coden; Boolean logic search method; and so on.

REFERENCE

(I) MILIS Appraisal Meeting Materials, Sponsored by Pao Sui-Loong Library of Shanghai Jiao Tong University

(2) Thompson, D.F., Campbell, M., An Introduction to MINISIS, International Development Research Centre, Ottawa, Canada, Jan. 1988

(3) BIBLIOFILE, ANY-BOOK System, The Library Corporation on CD-ROM, summer 1987

(4) Sally C. Tseng, CD-ROM Workshop, December, 1987

(5) A Multi-user and Online Library Circulation with Barcode Reader, International symposium on New Techniques and Applications in Library, Xian, China Sept.8-ll 1988.

(6) Report of Library Integrated System, International Symposium on New Techniques and Applications in Library, Xian, China. Sept 1988

NATIONAL UNIVERSITY OF SINGAPORE LIBRARY OUTLINES OF LIBRARY MANAGEMENT SYSTEM

ONG GIN HONG

1 Background Information

The National University of Singapore has eight faculties with 44 academic departments and three postgraduate schools. In addition, there are two special postgraduate and research institutions and other non-faculty teaching departments.

1.1 Library

The University's Library system consists of the Central, Chinese, Law, Medical, Science and Non Sui Sen Memorial Libraries. The total collection amounts to over 1.5 million books and about 14,000 current periodicals. It also holds approximately 29,000 reels of microfilms, 142,000 microfiches and over 4,000 titles of audiovisual programmes.

All major library operations are computerised (acquisitions, cataloguing, indexing and loans). Besides having access to local databases, the Library operates a Library Online Information Search Service (LOIS) which provides direct access to over 200 subject databases overseas.

1.2 Computer Centre

The Computer Centre provides a centralised computing facility to support the teaching and research functions of the University. The present academic computing facilities comprise an IBM 3081 KX2 mainframe computer system, approximately 2,000 micros and some 600 workstations/terminals distributed all over the campus. The Centre also supports the University Administration and Library by providing computer facilities and technical assistance to their respective computerisation programmes. The existing administration facilities comprise an IBM 3090/150 mainframe computer system (for office automation and decision support systems) and two NP 300/48.

The University is also linked to BITNET, a communications system that enables academics in the University to keep in touch with their colleagues in major universities and research institutes in the US, Europe, Japan and Australia to share information and databases with these universities.

1.3 Population of Students and Staff (1988/89)

Undergraduates:	14,317	
Postgraduates:	1,874	
	16,191	(total)
Academic Staff		
teaching:	1,301	
research:	157	
Admin Staff:	141	
Professional		
Library Staff:	57	
General Staff:	2,227	
	3,883	(total)

1.4 Proposed Campus Network

The University plans to put in place a campus-wide data network to support its educational and research requirements. The planned network will provide a fully integrated computing environment which will facilitate the access, sharing and optimisation of computing resources on campus. The planned network will also allow for access to public network services.

1.5 SILAS (Singapore Integrated Library Automated Services)

SILAS, the national online bibliographic network, was established in April 1987. It is based on a software developed by the Western Library Network (WLN), which runs on the IBM 3090 mainframe, maintained by the Computer Services Division (CSD), Ministry of Finance.

The NUS Library is one of the 31 participants of the SILAS bibliographic network, which serves government, academic and special libraries in Singapore. SILAS provides Singapore Libraries with cataloguing data and online access to bibliographic information through its huge pool of 4.59 million bibliographic records.

2 NUS Library Management System

The following areas have been identified for the NUS Library management system:

(a) Collection Development
(b) Personnel Management
(c) Budgetary Control
(d) Inventory Control
(e) Full Text Databases for:
  - Minutes of Senate Meetings
  - Minutes of Library Committee Meetings
  - Minutes of Important Project Meetings (i.e., Library Committee on System Upgrade)

Item (a) is already available in the library application systems (MINISIS and LACS) and Item (b) has also been provided by the NUS OA Decision Support System. Items (c)(e) will only be briefly discussed.

General Features of items (c) - (e):

- System is menu and command driven
- Moving from function to function is controlled by function keys or by user keyboard input
- System requires user logon identification and limits access to data bases by access control
- System provides for workforms or templates for entry/modification of data
- System must have powerful report generator and sophisticated word processing capability
- Only current information would be kept in online-active mode

2.1 Budgetary Control System

There will be 2 different types of databases in this system:

- 5-year Budget for New Equipment (YYYY/YYYY) e.g. (1988/1992)
- Annual Estimates (YYYY/YYYY) e.g. (1988/1989)

Data elements

- Department
- Item requested
    - Name of item
    - Cost per unit
    - No. of unit requested
    - Sum approved
    - Sum committed
    - Purchase Order No./Purchase Order Date
    - Invoice No.
    - Remarks

Search by Department/Purchase Order No./Purchase Order Date/ Invoice No./Invoice Date

2.2 Inventory Control

Data Elements

- Item Inventory Control No*
- Item Description*
- Item Series No*
- Purchase Order No*
- Purchase Order Date*
- Delivery Order No*
- Delivery Order Date*

NATIONAL UNIVERSITY OF SINGAPORE LIBRARY OUTLINES OF LIBRARY MANAGEMENT SYSTEM

ONG GIN HONG

1 Background Information

The National University of Singapore has eight faculties with 44 academic departments and three postgraduate schools. In addition, there are two special postgraduate and research institutions and other non-faculty teaching departments.

1.1 Library

The University's Library system consists of the Central, Chinese, Law, Medical, Science and Non Sui Sen Memorial Libraries. The total collection amounts to over 1.5 million books and about 14,000 current periodicals. It also holds approximately 29,000 reels of microfilms, 142,000 microfiches and over 4,000 titles of audiovisual programmes.

All major library operations are computerised (acquisitions, cataloguing, indexing and loans). Besides having access to local databases, the Library operates a Library Online Information Search Service (LOIS) which provides direct access to over 200 subject databases overseas.

1.2 Computer Centre

The Computer Centre provides a centralised computing facility to support the teaching and research functions of the University. The present academic computing facilities comprise an IBM 3081 KX2 mainframe computer system, approximately 2,000 micros and some 600 workstations/terminals distributed all over the campus. The Centre also supports the University Administration and Library by providing computer facilities and technical assistance to their respective computerisation programmes. The existing administration facilities comprise an IBM 3090/150 mainframe computer system (for office automation and decision support systems) and two NP 300/48.

The University is also linked to BITNET, a communications system that enables academics in the University to keep in touch with their colleagues in major universities and research institutes in the US, Europe, Japan and Australia to share information and databases with these universities.

1.3 Population of Students and Staff (1988/89)

Undergraduates:	14,317	
Postgraduates:	1,874	
	16,191	(total)
Academic Staff		
teaching:	1,301	
research:	157	
Admin Staff:	141	
Professional Library Staff:	57	
General Staff:	2,227	
	3,883	(total)

1.4 Proposed Campus Network

The University plans to put in place a campus-wide data network to support its educational and research requirements. The planned network will provide a fully integrated computing environment which will facilitate the access, sharing and optimisation of computing resources on campus. The planned network will also allow for access to public network services.

1.5 SILAS (Singapore Integrated Library Automated Services)

SILAS, the national online bibliographic network, was established in April 1987. It is based on a software developed by the Western Library Network (WLN), which runs on the IBM 3090 mainframe, maintained by the Computer Services Division (CSD), Ministry of Finance.

The NUS Library is one of the 31 participants of the SILAS bibliographic network, which serves government, academic and special libraries in Singapore. SILAS provides Singapore Libraries with cataloguing data and online access to bibliographic information through its huge pool of 4.59 million bibliographic records.

2 NUS Library Management System

The following areas have been identified for the NUS Library management system:

(a) Collection Development
(b) Personnel Management
(c) Budgetary Control
(d) Inventory Control
(e) Full Text Databases for:
  - Minutes of Senate Meetings
  - Minutes of Library Committee Meetings
  - Minutes of Important Project Meetings (i.e., Library Committee on System Upgrade)

Item (a) is already available in the library application systems (MINISIS and LACS) and Item (b) has also been provided by the NUS OA Decision Support System. Items (c)(e) will only be briefly discussed.

General Features of items (c) - (e):

- System is menu and command driven
- Moving from function to function is controlled by function keys or by user keyboard input
- System requires user logon identification and limits access to data bases by access control
- System provides for workforms or templates for entry/modification of data
- System must have powerful report generator and sophisticated word processing capability
- Only current information would be kept in online-active mode

2.1 Budgetary Control System

There will be 2 different types of databases in this system:

- 5-year Budget for New Equipment (YYYY/YYYY) e.g. (1988/1992)
- Annual Estimates (YYYY/YYYY) e.g. (1988/1989)

Data elements

- Department
- Item requested
    - Name of item
    - Cost per unit
    - No. of unit requested
    - Sum approved
    - Sum committed
    - Purchase Order No./Purchase Order Date
    - Invoice No.
    - Remarks

Search by Department/Purchase Order No./Purchase Order Date/ Invoice No./Invoice Date

2.2 Inventory Control

Data Elements

- Item Inventory Control No*
- Item Description*
- Item Series No*
- Purchase Order No*
- Purchase Order Date*
- Delivery Order No*
- Delivery Order Date*

- Cost**
- Supplier/Manufacturer*
- Location*
- Remarks

*Search keys

**Able to search through relational operators (, , = ,etc.)

2.3 Full Text Databases for Meetings Minutes

The Minutes of different Meetings would be kept in different files.

Data Elements

- Date/Time of Meeting*
- Venue of the Meeting
- Present/Absent
- Objectives of the Meeting
- Topics Discussed
- Headings**
- Follow-up action

*Search by date

**Search by key terms

MINISIS VERSION H APPLICATIONS

BOB McKERCHER
IDRC

1.0 INTRODUCTION

The evolution of the MINISIS system into an applications development language (fourth generation programming language) will put it into the same class as software packages such as dBASE IV, ORACLE and ADABAS. This programming language will make it possible for any MINISIS user, including many who have little or no experience with programming languages, to develop their own applications. Users will no longer have to resort to the use of programming languages such as SPL or COBOL if they are not satisfied with the MINISIS we now know. It may indeed become the norm for users to develop their own applications or to modify applications distributed by IDRC.

2.0 GENERALIZED APPLICATION SPECIFICATION (GAS)

The GAS or Standard MINISIS Application will be functionally compatible with and will serve as a replacement the version of MINISIS we now know. However, care should be taken not to form the impression that the GAS will be the main feature of MINISIS Version H. The most important feature of version H will be the Application Specification Toolbox. The GAS will only be one example of the use of that toolbox. Users may accept it as distributed, they may modify it, or they may build their own application, as they see fit.

2.1 HOW THE GAS WILL DIFFER FROM VERSION G

While the GAS will be functionally compatible with version G, it will include new features, many of which have come from our list of enhancement requests. Thus, we will see such things as user-defined word delimiters for record display, a greater number of sort keys, automatic garbage collection and more simplified inversion.

The biggest difference in the GAS will be in the interface and the structural implementation of its functions. Although functionally, it will duplicate version G, it will appear on the screen in a very different manner. Attached to this paper is a structure chart of the top level of the GAS which illustrates its functional decomposition. There will be two processors: one for the database manager, and the other for the end user. The database manager processor will be quite similar to what is in version G today, although it will include new features such as a form editor for creating data entry screens and a more streamlined inversion command. The end user processor however, will be constructed in such as way that users will not have to continually traverse menus in order to process data. The GAS top level structure chart shows how, on entry into the processor, the user opens a database, uses what we now know as QUERY to select records (or indicates that new records are to be created) and then goes on to process those records further by sorting, displaying or performing computations.

2.2 THE GAS DESIGN AND IMPLEMENTATION TEAM

The development of the Generalized Application Specification has just completed the Rough Functional Specification (RFS) stage. We are now entering the Detailed

Functional Specification (DFS) stage and anticipate this finalized functional spec to be completed sometime in mid-November. The GAS is being designed and will be implemented by the following people:

QUERY: May Ho
UPDATE and EXPORT: Mary Campbell
SORT: Terry Gavin
DBMGR and COMPUTE: Bob McKercher

This team will be assisted in its deliberations by Richard Lee (serving in the capacity of eminence grise), as well as members of the interface and documentation core groups. Some new features which these people have indicated they would like to include are:

SORT: the ability to review online the operation of the sort key specification to see whether it is performing as expected before submitting the entire sort job to MINISIS

EXPORT: a replacement for PRINT which will also incorporate ISOCONV

COMPUTE: the ability to write the result of a computational operation to a database field

If you have any questions or comments for these people, please speak to them during the conference. If you feel that you would like to make suggestions regarding their modules, feel free to contact them directly, but please be aware that no suggestions can be entertained after November 1, 1989 and also note that at this point, we cannot consider any requests which may require changes to the AST primitives, which are now in an advanced design phase.

3.0 INTEGRATED LIBRARY SYSTEM (ILS)

The ILS will probably not be distributed on the version H.00 release tape. Rather, it will most likely be included on a subsequent major (e.g. version I) or minor (e.g. H.01) release. Presently we are still very early in the design stages and will not be beginning the Rough Functional Specification until later this year.

3.1 SOURCES OF INPUT

The ILS functional spec will be produced in accordance with input received from the IDRC Library, the Shanghai Jiao Tang University, and several Canadian libraries which either now have or are considering the implementation of an integrated library system. The ILS design and implementation team will visit several sites before coming up with a core set of specifications. Integrated Library Systems are complex and many support a huge number of functions. We feel it is important at this stage to identify the most valuable of those functions for implementation while leaving other, more minor functions to perhaps later implementation or to implementation at user sites through an application extension.

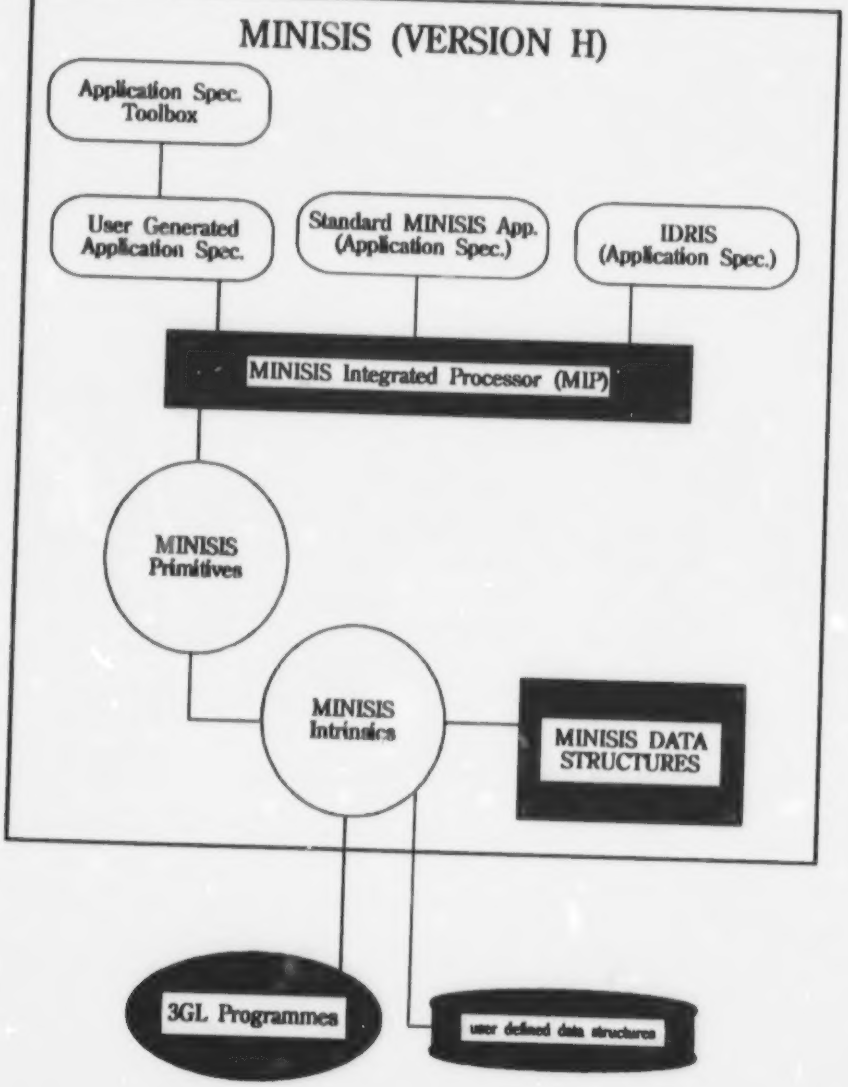
We are very keen to hear from MINISIS users who have comments about what functions they would like to see implemented. At present we have only the attached very tentative top level Rough Functional Specification. Some questions which this raises are: should we support a distinct ILL function at this stage? If so, should it be a separate module or perhaps part of the circulation module? Similarly, where should Serials Control be placed: stand alone or part of Acquisitions. What kind of interface should the OPAC have? Please let us know your views.

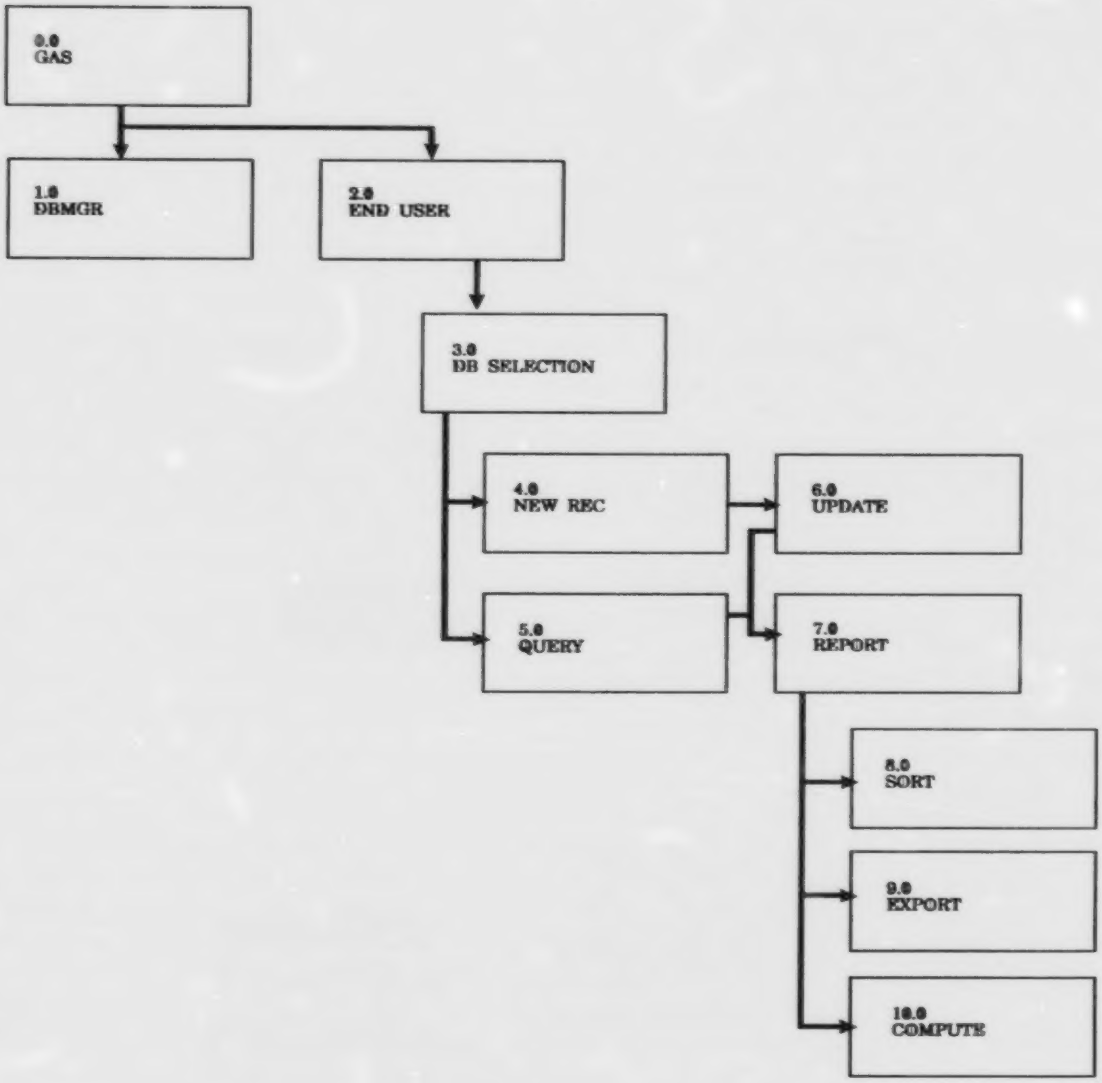
3.2 THE ILS DESIGN AND IMPLEMENTATION TEAM

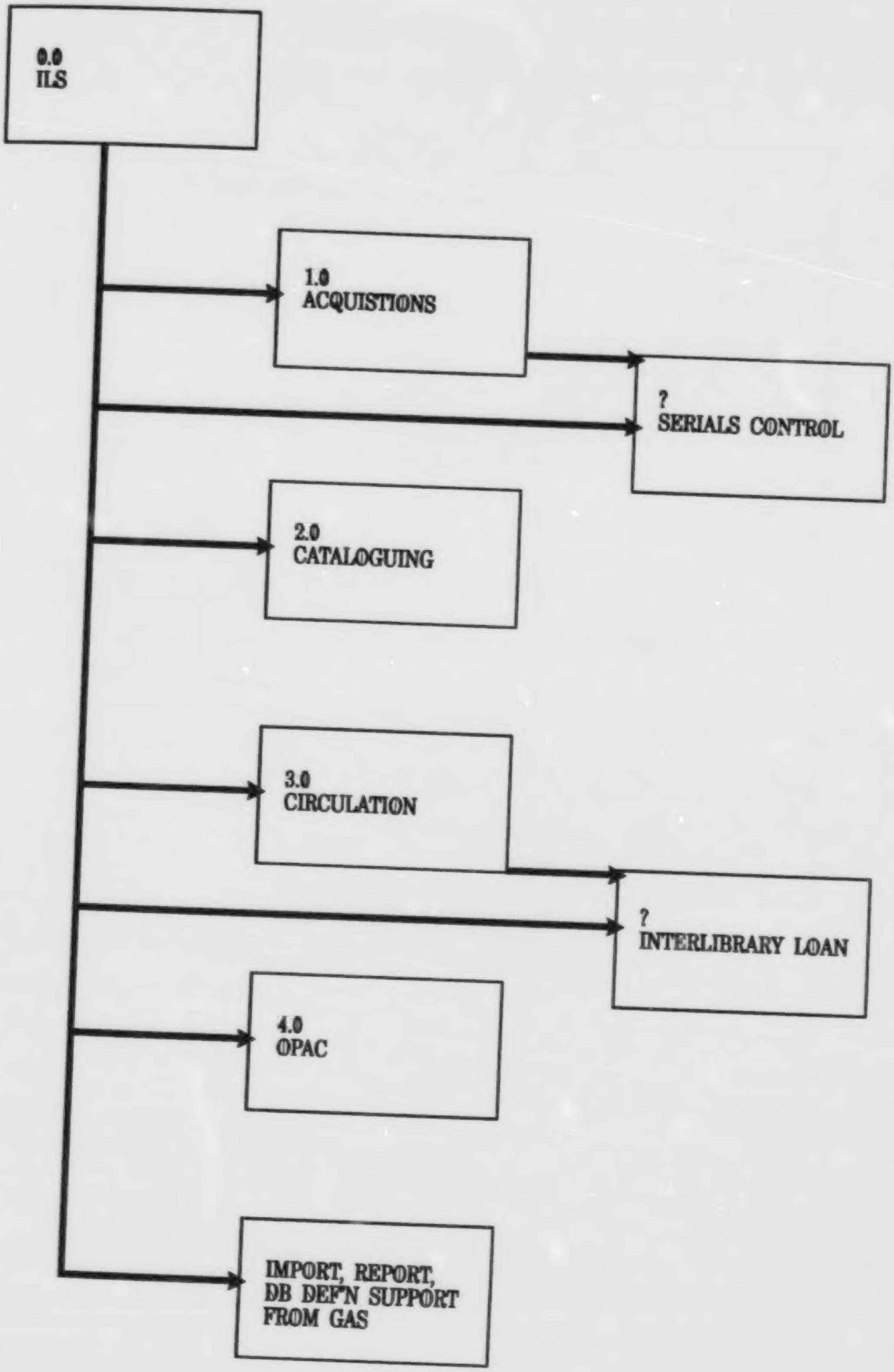
The ILS will be designed and implemented by May Ho, Yan Chun and Bob McKercher. Ms. Yan will be on loan to IDRC as an intern for two years from the Shanghai Jiao Tung University.

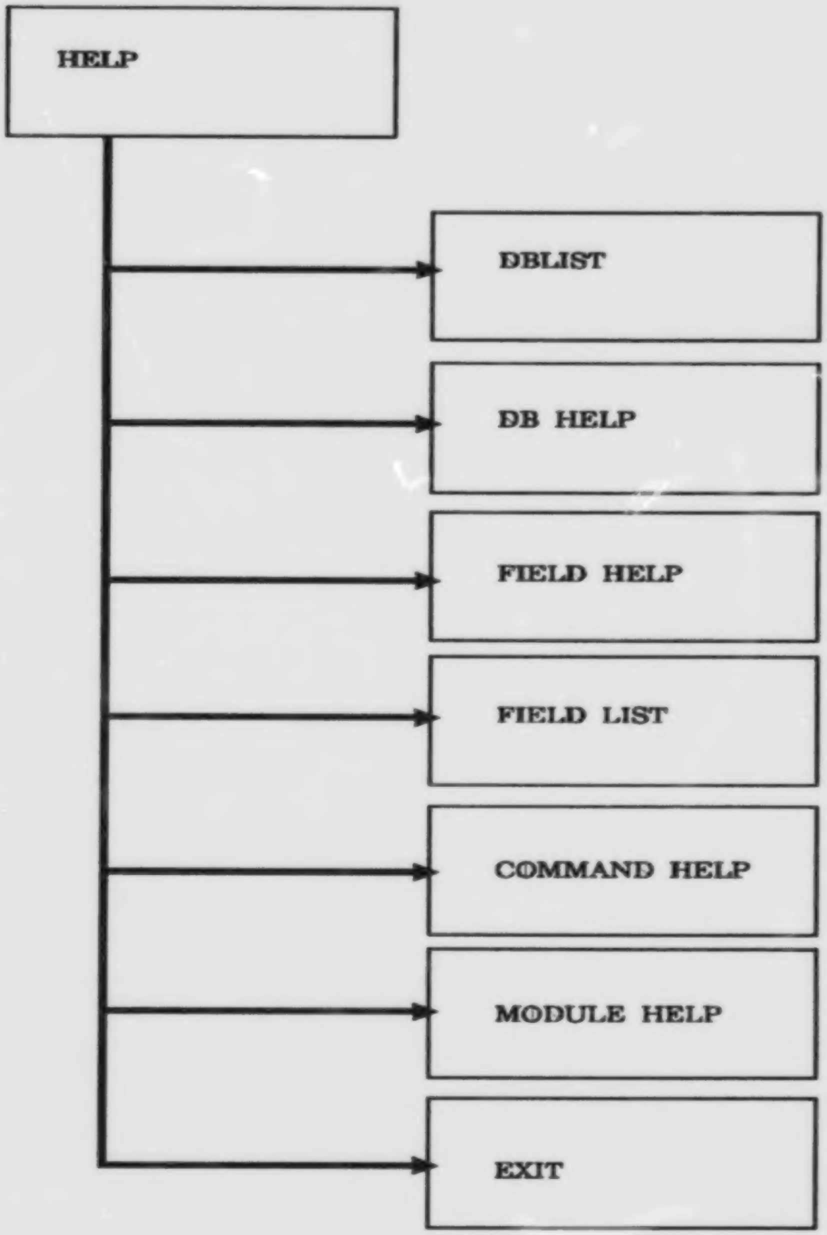
4.0 THE IDRIS APPLICATION

The IDRIS project management application will be a superset of the GAS; it will be designed and implemented by Mary Campbell and will be the first application of version H to be demonstrated. The IDRIS application will enter the DFS stage immediately once the GAS DFS has been finalized. Any comments or questions regarding IDRIS should be directed to Mary.


MINISIS (VERSION H)
Application Spec. Toolbox
User Generated Application Spec.
Standard MINISIS App. (Application Spec.)
IDRIS (Application Spec.)
MINISIS Integrated Processor (MIP)
MINISIS Primitives
MINISIS Intrinsics
MINISIS DATA STRUCTURES
3GL Programmes
user defined data structures


0.0
GAS
1.0
DBMGR
2.0
END USER
3.0
DB SELECTION
4.0
NEW REC
5.0
QUERY
6.0
UPDATE
7.0
REPORT
8.0
SORT
9.0
EXPORT
10.0
COMPUTE


0.0
ILS
1.0
ACQUISTIONS
?
SERIALS CONTROL
2.0
CATALOGUING
3.0
CIRCULATION
?
INTERLIBRARY LOAN
4.0
OPAC
IMPORT, REPORT,
DB DEF'N SUPPORT
FROM GAS


HELP
DBLIST
DB HELP
FIELD HELP
FIELD LIST
COMMAND HELP
MODULE HELP
EXIT

IDRIS - an Inter-Agency Development Research Information System

Mary Campbell
IDRIS Project Coordinator
International Development Research Centre

The objective of the Inter-Agency Development Research Information System (IDRIS) is to provide a common data base of information describing research activities located in or concerned with developing countries.

The contributors of this information are a group of donor agencies, including the IDRC (International Development Research Centre), which have the common goal of supporting technological research and development in developing countries.

Each contributor sends data to a data base which is located on the Hewlett-Packard 3000 computer belonging to the Information Sciences Division of IDRC. The data base is in MINISIS format.

Access to this data base is not restricted to the contributing agencies; any individual researcher or organization is welcome to search the IDRIS data base, or to ask one of the contributing agencies to perform a search. Other copies of the IDRIS data base are mounted at the following MINISIS installations:

Japan International Cooperation Agency (Tokyo)
International Labour Office (Geneva)
Agricultural University of the Netherlands
Arab League Documentation Centre
International Livestock Centre for Africa
Institute of Development Studies
Fondation du Roi Abdul-Aziz al Saoud (Morocco)
Gulf States Information Documentation Centre (Iraq)
Philippines Dept. of Science & Technology (Manila)

These copies are updated every 6 months.

History of the IDRIS data base

Before MINISIS was developed at IDRC, the Centre's data bases resided on an IBM mainframe computer at a local service bureau. The ISIS software system was used to maintain the Centre's data. Among the data bases on this system was the PINS (Project Information System) data base in which was stored information about all of the projects funded by the Centre. When MINISIS was first installed at the Centre in late 1977/early 1978, among the first data bases to be converted from ISIS format to MINISIS format was PINS.

PINS continued to be an important tool used by the Centre to provide information about past and current projects, both to program officers within the Centre and to scientists and policy makers in the research community as a whole.

IDRC belongs to an informal group of donor agencies known as the "like-minded agencies" because of the similarity of their objectives and mandates. This group also includes the donor agencies GATE (German Appropriate Technology Exchange); BOSTID (Board on Science and Technology for International Development) in Washington, D.C.); IFS (International Foundation for Science) and SAREC (Swedish Agency for Research Cooperation with Developing Countries), both situated in Stockholm; NUFFIC (Netherlands Universities Foundation for International Cooperation); and ACIAR (Australian Centre for International Agricultural Research). None of these agencies, other than IDRC, had an automated system of management for their project data. At meetings of this group, IDRC persuaded the heads of the other agencies that it would be a good idea to expand the PINS system to include information from the other contributing agencies.

The other members of the group agreed to take part in a pilot study to determine the feasibility of a common data base of research activities. IDRC provided the computer, the MINISIS system, and the person who would act as Data Base Manager for the IDRIS data base.

Structure of the IDRIS data base

It was realised that the PINS data base format, while adequate for the Centre's needs, did not meet all of the requirements of a data base which contained data from more than one contributing agency. While these agencies shared a common interest, the exact nature of their activities and the ways in which they administered the projects that they funded differed from organization to organization. At the same time, with data being entered into the data base from several sources, it was essential to establish certain standards so that agencies would be able to search for each other's data using the same criteria that they would use to search their own. In other words, the data base format had to be flexible enough to describe different kinds of research activities, while using the same format regardless of which activity was being carried out.

With this in mind, a committee formed of information officers from each contributing agency met to design the IDRIS data base. Information was described as being one of three kinds: essential (data must be present in the record); mandatory (data must be entered if the information is available); and optional (data is entered at the discretion of the contributing agency). Each field was assigned one of these categories.

The following standards were adopted: for codes of languages, countries and currencies, the International Standards Organization codes; for subject descriptors and names of countries, the Macrothesaurus for information processing in the field of economic and social development; and for names of institutions, IDRC's ACRONYMS data base.

A Systems specification manual was written to accompany the IDRIS worksheet, describing the more than fifty fields defined in the data base and defining rules for data entry into these fields. The fields include: the title, or name, of the project; donor; amount of grant; other contributing agencies; institute(s) at which the research takes place; duration of the research; names of researchers involved; titles of resulting publications; and subject descriptors and an abstract which describe the activity.

Because of the differences between the contributing agencies, it was agreed that no data base definition could possibly contain all of the fields required by each agency to store all of its information. Some agencies also preferred to keep certain information for their eyes only. The answer was to create "private" fields at the request of each agency, which would be visible only to that agency and the Data Base Manager. These fields are not defined on a worksheet or in a manual. An example of this type of private field is a field containing the date of birth of the researcher, for an agency which has as its mandate the support and encouragement of young scientists.

Problems encountered

Record size

In both the PINS and IDRIS data bases, one project occupies one record in the data base. The number of records in IDRIS is therefore much smaller than, for example, the number of records in a mailing list data base or in a data base which describes the holdings of a library. While there are only about 6000 records in IDRIS, each of these records describes an activity which may take as long as two or three years to complete. The result is a data base with relatively few but relatively large records.

The problem of size was a very real one since these records hold large amounts of textual data, including an abstract in one and sometimes in two languages. Under earlier versions of MINISIS, an individual record could only hold up to 4000 characters. The solution in PINS was to split the data base into two RD's: one to hold general information about the project, including title and abstract, and one to hold information about the research institute where the project took place, including the name(s) of the scientist(s) involved in the activity. The two RD's were then joined in a DS on the project number

While this solution was successful in PINS, it was rejected for IDRIS by the consultants who worked on the first phase of the activity. In the first place, although all of IDRC's projects had a number, other agencies did not assign a number to their projects and therefore their records could not be joined through this field. Secondly, the input operators who entered data for the other agencies were not familiar with MINISIS and found it difficult to switch from one data base to another during data entry. The solution for IDRIS was to allow input operators to continue entering data about a project in a second record, if the data was too large to fit into one record. The two records were then joined in a simple chain join. With the installation of MINISIS Version G, the chain join is no longer required; all data about a project is entered into a single record.

Security

The other major problem that was encountered was that of security. The IDRIS system is not unlike many other MINISIS applications in that the central data base resides in MINISIS format at an organization which acts as the central node in the information network. Other participating agencies send data to the central node. However, in most applications of this nature the data is sent in the form of worksheets or in machine-readable format, and is input at the central node, under the control of data base managers at this organization.

IDRC wanted to offer the IDRIS participants more options as far as data entry was concerned. While accepting data on worksheets or in machine-readable format, we also wanted to give the other agencies the opportunity to access our computer online through the Canadian Datapac telecommunications network, to enter data as well as to search for it. We were dealing with a group of agencies which, with the exception of IDRC, had no automated systems of their own. These agencies wanted to use the features of MINISIS for data input, sorting and printing; for example, two of the agencies use MINISIS to output directories of their own projects. The question that we had to answer was: how does an organization allow users from outside to enter data into a common data base, while still protecting the data base and the other applications on the system?

The answer was to use a combination of MINISIS and MPE security features. Each agency is given an individual user id and home group. User names and groupnames correspond to the acronym of the agency; e.g. SAREC, BOSTID, GATE etc. In each user's MINISIS security profile, the user is given "read" access to the IDRIS data base, and read/write access to another data base, a PS with the same name as the user id. The user's PS is used for data entry, and for searching, sorting and printing only that agency's data. The IDRIS data base is the common data base used for searching on all activities.

One of the IDRIS common fields is Polo, the Participating Agency acronym. This field is inverted and is used to search for activities funded by a particular donor agency. It is also used to control read access to the IDRIS records, through an initial restriction which refers to this field, in each of the individual PS's.

For example, the user BOSTID has a security profile, through which she is allowed read access to the IDRIS data base, and read/write access to a PS called BOSTID. The acronym for this agency in POIO is BOSTID, and the initial restriction on the BOSTID PS is the statement POIO BOSTID. This means that the user can only enter, modify and delete records with this acronym; in other words, she can only access her own records in UPDATE or DELETE. Of course, all of the users have their own passwords at the MPE level.

The IDRIS data base itself is a PS which contains only the common fields defined in the Systems Specification manual. The underlying RD, IDRISX, contains all fields defined for all agencies, both common and "private". The individual PS's contain definitions of the private fields for that agency.

The users' UDC's are set up with the NOBRK option, so that the users access the MINISIS menu, but cannot break and go to the MPE level. However, since the users will be creating print formats and INDEX, QUERY and COMPUTE hitfiles, they are expected to purge their files when they are finished. To this end, the MPE commands PURGE, LIST and RENAME have been added to the users' MINISIS menus.

Current status of the data base

By 1986, the founding agencies had decided that the establishment of the IDRIS system was feasible. Other agencies were invited to contribute data, and to search the data base. The PINS data base was converted to IDRIS format, using BATCHIN.

At present there are about 6300 records in IDRIS, contributed by the founding participant agencies (BOSTID, IDRC, SAREC, IFS, GATE), the United Nations University (Tokyo) and the Japan International Cooperation Agency (Tokyo). These records include closed and active projects.

User levels

There are three levels of IDRIS users: the management group (also known as the technical working group), the participants, and the searchers.

The management group is composed of the founding agencies: IDRC, BOSTID, SAREC, IFS and GATE (NUFFIC was also a founding agency, but opted out of the program before the pilot phase was completed). This group meets every two years to discuss any required changes to the system methodology.

The participant group is composed of those agencies which contribute data to the IDRIS data base. This can be any organization which is involved in research in a developing country and which intends to contribute data to the IDRIS data base on an ongoing basis. It need not necessarily be a donor agency, and may or may not reside in a developing country. At present, this group includes the members of the management group, as well as UNU, JICA, ACIAR (Australian Centre for International Agricultural Research) and NORAD (Norwegian Ministry for Development Cooperation). The latter two organizations have formally expressed their intention of contributing data to the data base, but we have not yet received their first contributions. The IDRIS Project Coordinator

disseminates news and information to this group through short newsletters, using the ENVOY electronic messaging system.

The searchers are organizations who have direct or indirect access to the IDRIS data base, but who don't contribute data. Direct online access is limited at present to academic and government institutions in Canada, and agencies in the international development community, such as IMF, World Bank and USAID.

However, any individual or agency can have indirect access to the data base, by approaching one of the participants with a request for a search.

Dissemination of information

The management group realises that online search access to the IDRC computer is not feasible for many interested organizations who want to become part of the searchers group. Some participating agencies, including IFS, SAREC and IDRC, have published

directories of their project data.

IDRC also makes IDRIS available on tape or diskette in ISO format to any organization who would like to mount the data base on their own system. For MINISIS users, IDRC will provide, if requested, the IDRIS data definition in ISO format, and a default format file. For micro-CDS/ISIS users, IDRC will provide the data base definition, worksheet, print format, FST, and data in ISO format.

Costs

IDRC assumes most of the costs of technical support, including telecommunications costs after the searcher or participant has accessed the Datapac network. The participant or searcher is responsible for all telecommunications costs until he/she reaches Datapac.

Terms and conditions of IDRIS Participation

We ask participants to indicate to IDRC in writing that they accept the following conditions:

1. The participating agency will determine which of its research activities will be input to the data base.

2. The participating agency understands that search access to the IDRIS data base is open to all.

3. The term "IDRIS data base" refers to the data in those fields described in System specification and data base design for an inter-agency development research information system. Private fields for the sole use of the participating agency can be defined on request; however, the IDRIS Project Coordinator will have access to these private fields.

4. The participating agency agrees to be identified as a contributor to the data base, in information disseminated about IDRIS.

5. The participating agency may search the data base on behalf of any other organization or person, but may not allow any other organization or person to log onto the IDRC host computer using the participating agency's user identification.

6. The participating agency agrees to follow the methodology described in System specification and data base design for an inter-agency development research system when contributing data to IDRIS.

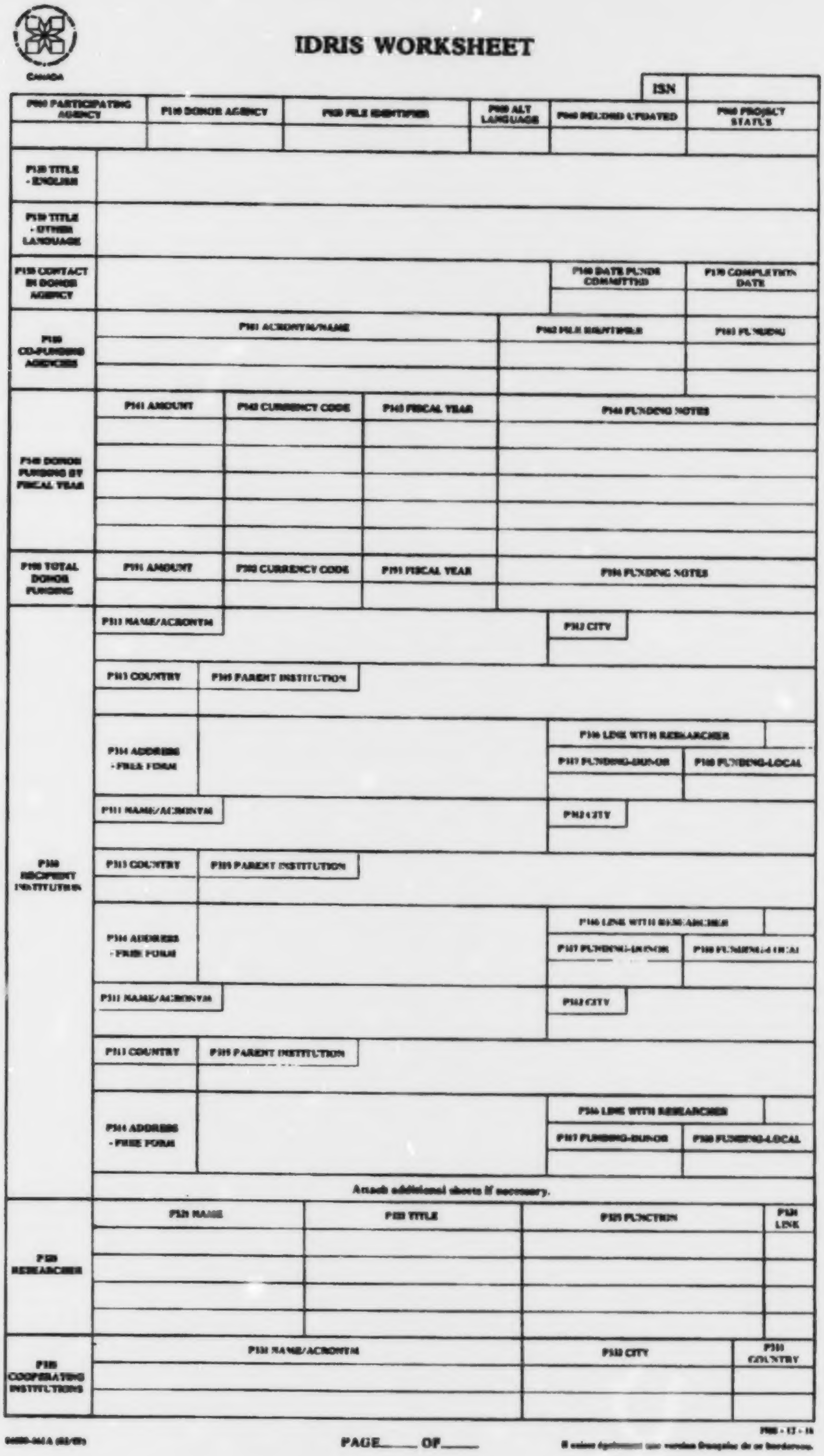
7. IDRC is responsible for technical support, including maintenance of the data base, software and host computer, access to the ENVOY electronic mailbox system as a means of communication, and dissemination of manuals and other guides such as the Macrothesaurus.

8. In the case of data entry by IDRC personnel from worksheets supplied by the participating agency, IDRC will charge the participating agency for the costs of data input.

9. The participating agency is responsible for the costs of reaching the Datapac network, for access to the IDRC computer and the ENVOY electronic mailbox; IDRC will be responsible for all other costs of telecommunications.

10. The IDRC Library will offer retrieval access to its data bases to the participating agency.

```
Wed, 4 Oct 1989                  LISTDOT DATA DEFINITION
                                                    O  L    N D  R S P M D B V P I I I I I N   T L   S S A   I    L V    A    E
                                                    F  E    U E  E U R A U I A R N N N N N U   Y E   T T N   N    O A    R    X
                                                    F  N    M C  P B O N P B L I V V V V V M   P N   R P T   V    G L    G    I
                                                    S  G    E    E F M D     I M                     I                        T
                                                    E  T    R P  A L P T C L D   O R O T T E   E E   P F F   F    C F    N
F I E L D   N A M E                   MNEMONIC TAG  T  H    I L  T D T O H V A K N E F Y H X   X X     I I   I    H I    A
                                                            C      E   R C L T E   L F P E T   T T   O L L   L    G L    M
                                                                   D   Y K   E Y       E S R   R R   P E E   E      E    E
..................................................................................................................................
*** Database : IDRIS

Participating Agency                    AGENCY P010 -1 10            Y           Y     B   1     10          P010
Agency File Identifier                  FILEID P020 -1 30        Y   Y           Y     B   1     30          P020
Alternate Language of Record            ATLANG P030 4  4             Y       Y                                      LANG V220
Last Date Record Updated                UPDATE P040 -1 10
Last Date Record Modified               MODIFY P050 -1 10
Status of project                       STATUS P060 -1 10            Y           Y     B   1     10          P060
Donor Agency                            DONOR  P110 -1 10            Y           Y     B   1     10          P110
Title -- English                        TITLE  P120 -1 160           Y           Y     B   99  W 40  N       P120
Title -- Other Language                 TITLEO P130 -1 250                       Y     B   99  W 40          P120
TOTAL FUNDING BY DONOR AGENCY           TOTALF P190 -1 100         Y Y
  Amount                                TOTAMT P191 -1 20
  Currency Code                         TOTCUR P192 -1 4
  Fiscal Year of Grant                  TOTFIS P193 -1 8
  Funding Notes                         TOTFUN P194 -1 50
DONOR FUNDING BY FISCAL YEAR            FUNDNG P140 -1 71        Y Y Y
  Amount                                AMOUNT P141 -1 10
  Currency Code                         CURNCY P142 8  4                     Y                                      ISOC V140
  Fiscal Year of Grant                  FISCYR P143 -1 8
  Funding Notes                         FNNOTE P144 -1 50
Contact in Donor Agency                 CONTAC P150 -1 80        Y   Y
Date Funds Committed                    FNDATE P160 -1 8             Y           Y     B   1     8           P160
Completion Date of Activity             FINISH P170 -1 8             Y
CO-FUNDING AGENCIES                     COAGEN P180 -1 135       Y Y Y
  Co-funding Agency Acronym/Name        CONAME P181 -1 100
  Co-funding Agency File Id.            COFILE P182 -1 25
  Funding from Co-funding Agency        COFUND P183 -1 15
RECIPIENT INSTITUTION                   INSTUT P310 -1 999       Y Y Y
  Institution Name                      INNAME P311 -1 200                       Y     B   99  W 40          P311
  Institution City                      INCITY P312 -1 40
  Institution Country Code              INCTRY P313 4  2             Y       Y   Y     B   1     50     AGEO P313   ISOC V110
  Institution Address -- Free Form      INADDR P314 -1 550
  Parent Institution                    PARENT P315 -1 100                       Y     B   99  W 40          P311
  Link with Researcher                  INLINK P316 -1 2
  Funding -- Donor Agency               FDONOR P317 -1 15
  Funding -- Local Contribution         FLOCAL P318 -1 15
RESEARCHER                              RESRCH P320 -1 1050      Y Y Y
  Researcher's Name                     RENAME P321 -1 50                        Y     B   99  W 40          P321
  Researcher's Title                    RETITL P322 -1 10
  Researcher's Function                 REFUNC P323 -1 30
  Link with Institution                 RELINK P324 -1 1
COOPERATING INSTITUTION                 COOPIN P330 -1 242       Y Y Y
  Cooperating Institution Name          CPNAME P331 -1 200                       Y     B   99  W 40          P331
  Cooperating Institution City          CPCITY P332 -1 40
  Cooperating Institution Country       CPCTRY P333 4  2                     Y                                      ISOC V110
                                                              ........................................................
Availability of Documents               AVLDOC P340 -1 150       Y   Y
Geographical Area under Study           AREA   P410 -1 200           Y           Y     B   99  T 50     ANYG P410
Macrothesaurus Subject Descriptors      THES   P420 4  250           Y           Y     K Y 99  T 50          THES
Non-thesaurus Subject Descriptors       NOTHES P430 -1 200           Y           Y     B   99  T 50          P430
Abstract -- English                     ABSTR  P440 -1 1300          Y
Abstract -- Other Language              ABSTRO P450 -1 1500          Y
Notes                                   NOTES  P510 -1 250           Y
ISO Country Name                        VLCTRY V110 36 50
ISO Currency Name                       VLCNCY V140 12 74        Y
UNISIST Language Code                   LANGC  V210 4  4
UNISIST Language Name                   LANGN  V220 10 50
Resequence number (for RENUM)           RESEQN R400 -1 10
  IDRC Division Acronym                 DIV    N020 -1 30                          Y B   99  W 40          N020
  IDRC Sub-division                     SDIV   N030 -1 200                         Y B   99  W 40          N020
  Sigle de la division (CRDI)           DIVF   F020 -1 100                         Y B   99  W 40          N020
  Sous-division (CRDI)                  SDIVF  F030 -1 200                         Y B   99  W 40          N020
  Region de la recherche                REGION F410 -1 200                         Y B   99  T 50  D       F610
  Descriptions macro des sujets         THESF  F420 -1 500
  Descriptions non-Macro                NOTHEF F430 -1 200
```



IDRIS WORKSHEET

ISN

Attach additional sheets if necessary.

PAGE____ OF____

P390 AVAILABILITY OF DOCUMENTS

P400 GEOGRAPHICAL AREA UNDER STUDY

P420 MACROTHESAURUS SUBJECT DESCRIPTORS

P430 NON-THESAURUS SUBJECT DESCRIPTORS

P440 ABSTRACT-ENGLISH

P450 ABSTRACT-OTHER LANGUAGE

P500 NOTES

IMPLEMENT SYSTEM OF ULTRASONIC DIAGNOSIS MANAGEMENT ON THE MINISIS DBMS

Li Shilang
Yao HAO
Wang Zhe
Computer Room Library
China-Japan Friendship hospital, Beijing,China

ABSTRACT:

The Management System of the Ultrasonic Diagnosis department on the MINISIS DBMS is applied firstly in our country. It plays an important role in every field of clinical medicine, especially in clinical therapy and medical research. This paper is divided into 4 sections, which describes the feature of ultrasound diagnosis, and describe Systematic research and analysis, the selected DB model, and the application and our prospective imagination of the system in detail.

INTRODUCTION:

The diagnostic application of ultrasound has been developing very rapidly in the world and has become an indispensable diagnostic modality in every field of clinical medicine. It is a diagnostic method which is connected with radar and the principle of acoustics, and is unimpaired, simple, easy.

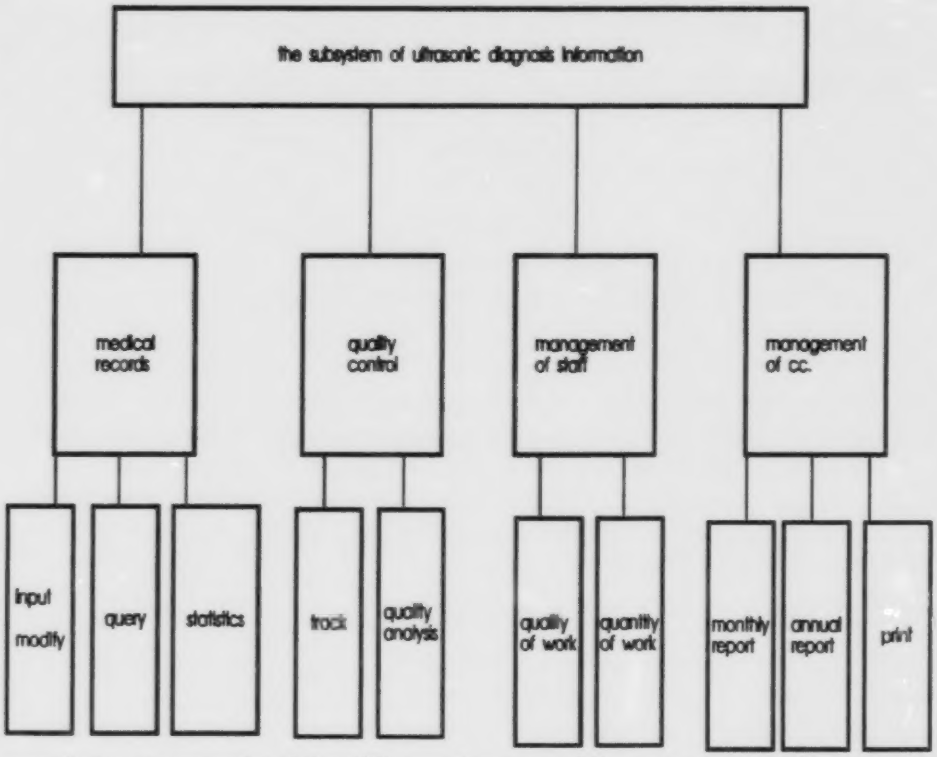
It is applied extensively in various clinical fields, especially, it is the best diagnostic method for soft tissues. It can explore organ or mass size, shape, thickness, depth and can differentiate between the physical features of the tissues, examination function and rules of activity for heart and vessels, especially, brain, heart, liver, bile, Kidney, gynaecologic disease, hydrothorax etc..

Our China-Japan Friendship hospital is famous in the world, we use a lot of advanced equipment. Our Department of Ultrasound Diagnosis is the largest Ultrasonic Diagnosis Department in our country, there are 20 units of ultrasonic diagnosis equipments, such as SDD-880,SSH-90A5,etc.. We have 30 doctors and nurses. We have collected a large quantity of cases of ultrasound therapy, it is valuable information of clinical diagnosis and medicine research. We have design and created the first Ultrasound diagnosis information database in our country. It is a large, useful, and complete system which can be used in clinical research, medical records, retrieval of disease, filing of clinical materials, etc.. It is very convenient for management of medical information.

I. Systematic research and analysis

We divided the system into four function modules, used the method of system engineering and based on the system analysis, according to user's demand as shown in the following chart:

Feature users:


the subsystem of ultrasonic diagnosis information
medical records
quality control
management of staff
management of cc.
input modify
query
statistics
track
quality analysis
quality of work
quantity of work
monthly report
annual report
print

1. The patients are examined for different items, and for one or more than one item, so it is the best to use variable-length records.
2. Some items need to be examined rapidly so they should be designed into the subfield.
3. It is operation of algebraic, and statistical quantity is little.
4. They used to examine about 70 patients every day on average, in total there are numerous members every year, so examination speed was requested to be quite fast.
5. They request more complex searches which can be carried out using more than one key and more than one inverted field, with the AND, OR, EOR, and NOT, ADJ, IGN expressions.
6. We compared with SQL, IMAGE and MINISIS at the HP-3000/70, and decided that we can use MINISIS.

Data base management system	MINISIS	IMAGE	SQL
type of dbms	relational DBMS	network DBMS	relational DBMS
application and language	VPLUS COBOL PASCAL SPL	VPLUS DICT-INF-REP TRANSACT COBOL FORTRAN PASCAL BASIC	VPLUS COBOL PASCAL
fixed or variable length	variable	fixed	fixed
repeatable fields (y/n)	y	n	n
subfields (y/n)	y	n	n
speed of search	fast	fast	not fast
computational functions	algebraic calculation	uses special language programming	uses special language programming
character sets	multi-character set capability	English and Chinese character sets	English and Chinese character sets
condition of application	without any programming	use TRANSACT language (simple and easy)	use COBOL or PASCAL programming languages
range of exchange information used to communicate with other DBs	can exchange data with dBASE III and IMAGE	can exchange data with MINISIS and SQL through an interface	can exchange data with MINISIS and other SQL applications through an interface

II. Select DB model

We made an detailed investigation on the work procedure of ultrasonic Diagnosis disease and management. We designed the work flowchart, data flowchart and made the data dictionary. Then, we created two databases, the CSZDK and CSZDK1. The CSZDK is the main database, it is an RD type, and the data structure is as following.

```
* 01 * ----------  数 据 库 头 标 屏 幕  ------------ * 01 *
数据库名称 ------ [CSZDK ] 数据库类型 ---------[RD]
                           数据模式 a(a,d)* ---[CSMS  ]
* 02 * 文件类型 ( 1=MASTER * 2=KSAM * 3=ANY * 4=THES)---[1]
非 MINISIS 文件类型   ( 1-32767 )  -----------[    ]

* 03 * 组名  --- [PUB     ]
MASTER 文件名 --  [MCDZDK  ] MASTER 文件最大容量 -- [100 ]
XRFF 文件名 ---   [XCSZDK  ] ISNS 最大值 ---- [10000 ]
自动编号 (Y/N) -----     [Y] 控制更新的 PS 数 (0-8) - [0]
用户定义的提示 (Y/N) --  [Y] 文献定位指示符 (Y/N) -- [ ]
自动删除文件中记存数(Y/N)·[ ] 自动删除文件名 --  [        ]
是否要联机MPE LOG 文件(Y/N)[ ] MPE LOG 文件名 --  [        ]
是否要联机倒排日志(Y/N) - [ ] 隐含检索字段标识符 --  [A200]
合法性检验用出口名称 ----  [     ]
隐含打印格式文件名 -----   [FCSZD      ]
参数文件名 ---------       [CCSZDK     ]
```

FIELD NAME: B CHAO RI QI
MNEMONIC: BCRQ
FIELD TAG: A100
MAXIMUM FIELD LENGTH: 10
PROMPT: Y
INVERSION: Y
INVERT ONLINE: Y
INVERTED FILE TYPE: B
NUMBER OF SIGNIFICANT KEYS EXTRACTED: 1
LENGTH OF STRING TO EXTRACT: 10
ROOT NAME OF INVERTED FILE: GBCR

We have designed the record to include the medical record number, date of examination, sex, age, code of examination region, result of examination, etc.. The doctors can rapidly search the region of examination, the result of examination of patients according to the number of medical records and date. They can also search the result of examinations looking for one in which somebody suffers from during a period. While making the data dictionary, we implemented a coding system using an internationally standardized code to make international exchange convenient, for example, name of disease, region of examination and the numbers of work which represent the staffs and so on. etc..

III. Condition of Application

After the DB had been created, the user could input more than a thousand medical cases, whenever they need, the doctors of every department can use ONLINE SEARCH conveniently to count the results of examinations and the quantity and quality of the work of every examiner, and economic activity of this department and so on. After that, the doctors and nurses can get rid of their notes, cards. The computerized information help them to make greater progress in medical information management. Figures 4 and 5 represent a record input and a monthly report.

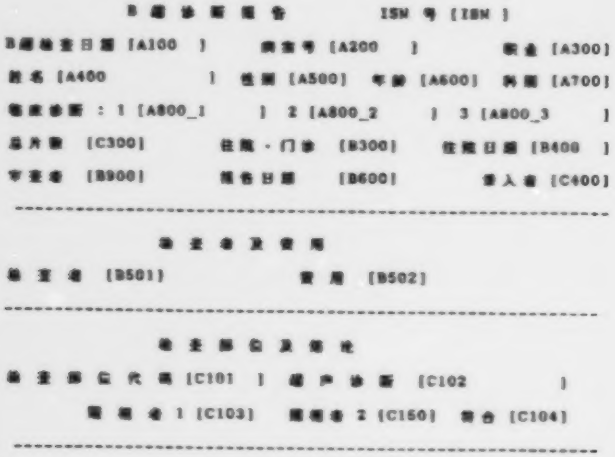


Figure 4: Data Entry

```
-----------------------------------------------------------

  超  声  诊  断  科  月  工  作  情  况  表     TUE, JUL 11, 1989

-----------------------------------------------------------

人员编号          收    入             人员编号          收    入

  总计            200

   1               10                     16              20

   2               20                     17              10

   3               30                     18              20

-----------------------------------------------------------
```

Figure 5: Monthly Report

IV. Prospective imagination of the system

The effort to use the ultrasonic equipment to diagnose disease and the MINISIS DBMS to manage the information management in the hospital is satisfactory. Therefore, we have decided to extend the application in other departments of the hospital, for example, various medical record, X-RAY diagnosis etc.. in order to manage medical information more efficiently.

Unfortunately, MINISIS compute function was not satisfactory for the compute, print function couldn't be simultaneously completed. We hope to connect with DBASEIII in order to build interactive communication of information between MINISIS and DBASEIII that will be satisfactory for the users.

We thank Mr. Shi Jianmin, director of Department of Ultrasonic Diagnosis for his honest support.

超声诊断管理系统在MINISIS数据库上的应用

中日友好医院 计算机室 李世良 杨好
图书馆 王劼

［摘要］：超声诊断管理系统的建立和应用，在我国超声诊断领域中尚属首创，它在临床治疗和医学研究中都起着重要作用。本系统是在MINISIS的基处上，在HP3000-70上实现的。本文首先叙述了超声诊断的特点，然后，对系统调査和分析、确定数据库模式、系统的应用情况、以及今后的设想四个部分都作了详细的说明。

超声诊断在医院的医疗诊断中是必不可少的。它是将雷达技术与声学原理结合起来，应用于临床医学的一种方法。

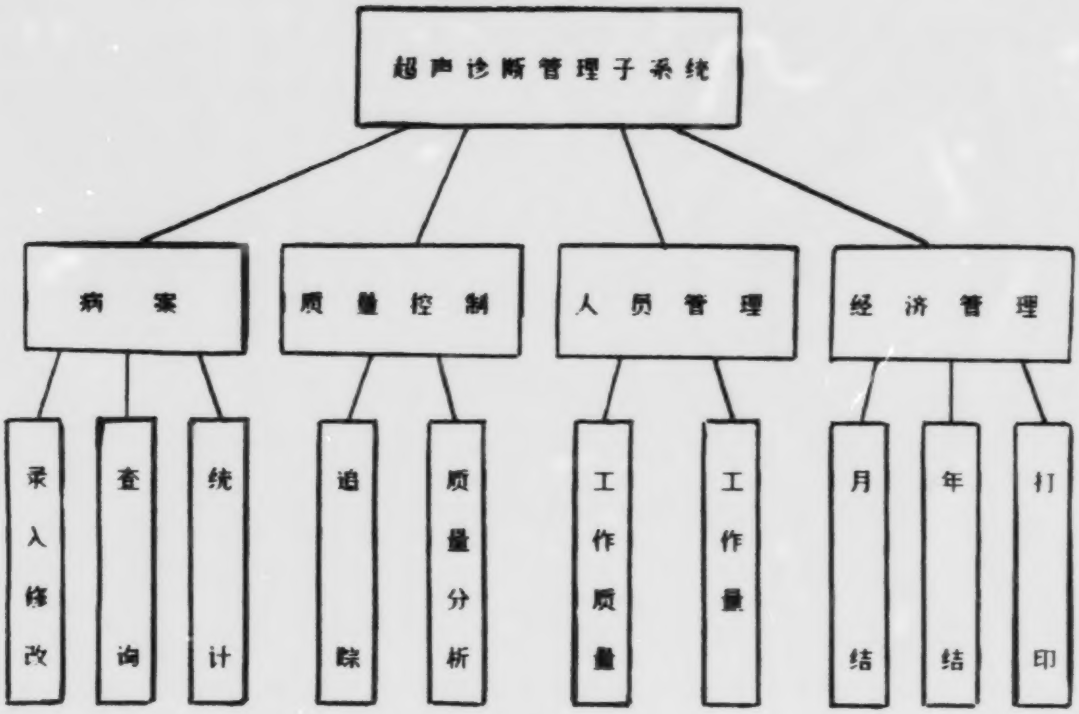
由于超声诊断具有无损伤的特点，特别对软组织的鉴别，充分显示其特长，受到人们的普遍重视。它可以探测到脏器或肿块的大小，形状，厚度，深度；鉴别组织的物理特性以及检测心脏和血管的功能和活动规律。尤其是在脑，心脏，肝，胆，肾，妇科病和胸腹水等方面的运用尤为广泛。

中日友好医院是全国著名的医院，它拥有全国最大的超声诊断科和先进的超声诊断设备近20台，例如：SAL-77AB型超声诊断仪，SDD-880彩色多普勒超声诊断仪，SSH-90AB型超声诊断仪，科内工作人员近30名。

我院接待的患者多为疑难病患者。大量积累的超声诊断病案是进行临床治疗和医学研究的重要资料。

一．系统调査和分析

根据系统工程的方法，我们在系统分析的基础上，结合用户目标和要求，将它划分为四个工能模块，如图1：


超声诊断管理子系统
病案
质量控制
人员管理
经济管理
录入修改
查询
统计
追踪
质量分析
工作质量
工作量
月结
年结
打印

图一．超声诊断科子系统功能图

用户的特点可归纳为以下六点：

１．不同患者检查的项目差异很大，从一，两项到十几项不等，因而最好采用可变长记录类型。

２．有的检查项目需要重复多次，且包含有子字段。

３．计算量不大，全部为代數运算。

４．每日检查人次约为７０人次，累计一年后，數目可观，因此要求查询速度快。

５．要求能作多种组合查询，对检查关键字可左，右截段。根据这些要求和特点，我们对在ＨＰ３０００／７０可运行的数据库ＳＱＬ．ＩＭＡＧＥ和ＭＩＮＩＳＩＳ作了如下比较（如图２）最后确定采用ＭＩＮＩＳＩＳ DBMS。

数据库名称	MINISIS	IMAGE	SQL
数据库类型	关系	网状	关系
支持软件及语言	VPLUS COBOL PASCAL SPL	VPLUS DICT-INF-REP TRANSACT COBOL FORTRAN PASCAL BASIC	VPLUS COBOL PASCAL
记录定长或不定长	不定长	定长	定长或不定长均可，但不定长记录效果不显著
是否可有重复字段	是	否	否
是否可带子字段	是	否	否
查询速度	快	快	较慢
计算功能	代数运算的功能较弱	由采用的语言和编程决定	由采用的语言和编程决定
文字种类	九种文字，包括中文和英文	中，英两种文字	中，英两种文字
应用情况	不用编程，可快速开发使用	需编程，采用TRANSACT可使程序编制工作量减少	需编程，且工作量较大
与其它数据库进行信息交换方式	可与DBASE3，CDS/ISIS等数据通讯	通过接口，与MINISIS,SQL进行数据通讯	通过接口，与IMAGE进行数据通讯

图 2．MINISIS，IMAGE，SQL 比较

二. 确定数据库模式

我们对超声诊断科的工作，作了详细的调查，画出了业务流程图，编制了数据字典(略)。然后确定了两个数据库， CSZDK和CSZDK1 。其中，CSZDK 是主要的数据库，它是RD类型，数据结构如图 3:

```
* 01 * ----------   数 据 库 头 标 屏 幕   ------------ * 01 *
数据库名称 …………… [CSZDK ] 数据库类型 ………………… [RD]
                          数据模式 a(a,d)* …… [CSMS  ]
* 02 * 文件类型 ( 1=MASTER * 2=KSAM * 3=ANY * 4=THES) …… [1]
非 MINISIS 文件类型  ( 1-32767 )  ……………………… [    ]

* 03 * 组名  ……… [PUB      ]
MASTER 文件名 ……  [MCDZDK  ] MASTER 文件最大容量 … [100 ]
XRFF 文件名 ………  [XCSZDK  ] ISNS 最大值 ………… [10000 ]
自动编号 (Y/N) ……………  [Y] 控制更新的 PS 数 (0-8) … [0]
用户定义的提示 (Y/N) ……  [Y] 文献级位指示符 (Y/N) …… [ ]
自动跟踪文件中记存取(Y/N)·[ ] 自动跟踪文件名 ……  [      ]
是否要联机MPE LOG 文件(Y/N)[ ] MPE LOG 文件名 ……  [      ]
是否要脱机倒排日志(Y/N) … [ ] 隐含检索字段标识符 …  [A200]
合法性检验用出口名称 ………… [     ]
隐含打印表格文件名 …………… [FCSZD      ]
参数文件名 ……………………… [CCSZDK     ]
```

接上图：

```
FIELD NAME : B CHAO RI QI
MNEMONIC : BCRQ
FIELD TAG : A100
MAXIMUN FIELD LENGTH : 10
PROMPT : Y
INVERSION : Y
INVERT ONLINE : Y
INVERTED FILE TYPE : B
NUMBER OF SIGNIFICANT KEYS EXTRACTED : 1
LENGTH OF STRING TO EXTRACT : 10
ROOT NAME OF INVERTED FILE : GBCR
  .
  .
  .
  .
  .
```

图 3．数据结构

其中，病案号，B 超检查日期，性别，年龄，检查部位代码，超声诊断结论等等作为倒排字段。这样医生可根据病案号和检查日期迅速找到他所需要的病人的检查结果。也可以统计出某段时间内患某种疾病的病人的检查结果。在编制数据字典时，字段名如疾病名称，检查部位等均采用国际标准编码，工作人员采用工号代替等等，这样采集的资料可以参加国际交流。

三．应用情况

数据库建成后，超声诊断科已输入了上千份病案，各科医生均可以方便地联机查询任何时间患有某种疾病的病人检查结果，科主任可以迅速地统计出在某段时间内本科工作人员的工作情况，工作质量的好坏，本科的经济情况等等。从此，医生和护士摔掉了登记本，卡片，从繁琐的登记查询统计工作中解脱出来，将精力更好地集中于检查患者疾病的工作中去。使医院的宗旨—患者至上，得以更好地贯彻。

下面是录入画面和月统计报表：

```
                    B超诊断报告                   ISN号 [ISN ]
B超检查日期 [A100   ]     病案号 [A200    ]            职业 [A300]
姓名 [A400               ] 性别 [A500] 年龄 [A600] 科别 [A700]
临床诊断 : 1 [A800_1        ] 2 [A800_2        ] 3 [A800_3        ]
总片数   [C300]         住院·门诊   [B300]    住院日期 [B400   ]
审查者   [B900]         报告日期    [B600]          录入者 [C400]
-----------------------------------------------------------------

                    检查者及费用

       检查者  [B501]                  费用  [B502]

-----------------------------------------------------------------

                    检查部位及结论

检查部位代码 [C101  ]  超声诊断  [C102              ]
             照相者 1 [C103]  照相者 2 [C150]  符合 [C104]
-----------------------------------------------------------------
```

图 4．录入画面

接上图：

```
-----------------------------------------------------------------
       超声诊断科月工作情况表      TUE, JUL 11, 1989
-----------------------------------------------------------------
人员编号          收    入          人员编号          收    入
  总计           10500
   1               310                 16              200
   2               120                 17              500
   3               250                 18              250
   .                                   .
   .                                   .
   .                                   .
   .                                   .
   .                                   .
-----------------------------------------------------------------
```

图5. 月统计报表

四. 今后的设想

超声诊断科采用MINISIS进行管理，仅仅是MINISIS应用于医院信息管理的一个尝试，从目前使用情况来看，用户比较满意，所以决定逐步推广到各科病案管理中去。例如：检验科，X光科科等等。

令人遗憾的是，MINISIS的计算功能较弱，统计，计算，打印不能同时作，所以想将它与DBAS3结合起来使用，建立MINISIS与DBASE3之间交互式的通迅方式，以便更好地满足用户的要求。

在本系统的开发过程中，中日友好医院超声科主任石健民先生给与了热情支持和大力协助，我们对此表示感谢。

SELECTIVE DISSEMINATION OF INFORMATION USING MINISIS AT THE NATIONAL INFORMATION CENTRE, S.N.D.T. WOMEN'S UNIVERSITY, BOMBAY

Ms. Harsha Parekh (Deputy Librarian - Co-ordinator)

The National Information Centre for Sociology, Women's Sturdies, Home Science, Special Education, Library Science and Gujarati at the Shreemati Nathibai Damodar Thackersey (SNDT) Women's University, Bombay provides a selective dissemination of information service using a locally created database. The establishment and maintenance of the database and the SDI service are undertaken using the MINISIS software. This paper describes the rationale of the service and highlights some of the issues involved in managing the service.

THE RATIONALE

The Centre was set up in 1986 by the University Grants Commission (India) as part of its larger efforts to improve teaching and research in institutions of higher learning by enhancing support services to scholars. The Centre was expected to provide services such as selective dissemination of information, literature searches, referral, document delivery, etc. which would improve both indicative access to information and physical access to documents.

Selective dissemination of information – i.e. channeling information about new documents selectively – was expected to help scholars identify current materials specific to their interest and thus face the "information explosion" more systematically and confidently. The SDI service is usually based on computerized bibliographic databases and is provided by many information vendors.

For most of the developing countries, including India, online access to these vendors has been well nigh impossible because of high costs and unreliability of telecommunication facilities. The spread and easy access to computer facilities in the country during the 1980s has made possible the alternative of acquiring tapes of bibliographic databases and providing an in-house SDI service based on them. Such a strategy has been adopted by a number of institutions in India e.g. the National Science Information Centre at the Institute of Science (Bangalore), Central Food Technology Research Centre (Mysore), and the Bhabha Atomic Research Centre (Bombay).

While in some subject area, such as the natural sciences and technology, this strategy is extremely useful, it does not seem to be the best way to deal with other subject areas like the social sciences. The provision of SDI in such a manner assumes that the bibliographic databases acquired do contain a satisfactory amount of published literature which is both relevant and timely to meet the information needs of users. This assumption does not hold good for many fields.

The areas of interest to the Centre were varied, ranging from the humanities and social sciences to multidisciplinary mission oriented fields. An initial analysis revealed the varied orientation of the subjects, their literature and bibliographic services. As the Table 1 indicates, international computerized databases either did not exist or did not include sufficient relevant material. An examination of the Sociological Abstracts for a six month period revealed that only 101 items from 7 journal titles from India were included. Of these about 70 were more than 18 months old. It was estimated that during the period approximately 2000 relevant items had been published in the country.

The analysis also indicated that information needs and literature use patterns in many of the fields were culture specific; scholars were frequently interested in local issues and required indigenous literature for facts, data and discussions on problems. International literature was used chiefly to understand theories, define concepts, evolve methodologies and make comparisons.

As such, it was decided to create and maintain a database emphasising Indian materials and provide SDI services based on it.

THE SERVICE

The MINISIS software was acquired in September 1987 and after a month trial period, the database was set up. Data input work which began in November 1987 was initially restricted to Women's Studies, Sociology and Home Science. Beginning from the middle of 1988, a quarterly SDI service has been offered. Four altering notifications in the three subject areas have been sent out by the middle of 1989. The first service went out to 100 users. By the fourth bulletin the number had grown to 300 users. This section describes some of the relevant structural characteristics of the three databases comprising the data submodel and outlines the activities and procedures followed.

The Bibliographic Database: SUCHAK

SUCHAK – literally meaning "one that informs" – is a large database covering all subject areas of the Centre's interest. It emphasizes Indian materials and includes references to articles in more than 400 journals (including popular magazines and newspapers), books, theses, dissertations, conference papers, reports, etc. since 1987. Older material is selectively being added. At present restricted to materials in English. Hindi, Marathi and Gujarati, it is proposed to cover other Indian languages in the future.

The database includes references available in a number of libraries. While materials available in the S.N.D.T. Women's University system form the major portion, some important additional data obtained from other libraries, either by visits of the Centre's staff or voluntarily sent in by the cooperating institutions. Since there is a considerable degree of overlap between the different subjects, SUCHAK is set up as one large database with projected subsets for each subject.

In designing SUCHAK, the UNISIST Reference Manual for Machine Readable Bibliographic Description was used as a guideline. Simplification was achieved by limiting the number of fields, selecting a limited number of subfields within each field and by making some fields not repeatable. For describing the subject contents, thesauri based controlled vocabularies have been used. This is supplemented whenever necessary by uncontrolled terms.

The author, subject, language, origin of item (Indian or foreign), geographic descriptor, target audience (i.e. research level or general) fields are inverted. The inversion is on release.

Since SUCHAK is used both for retrospective searches and SDI, it is necessary to distinguish the subset on which SDI is to be run. A 'status' field is used for this purpose.

Records to be included in the next SDI are marked 'Current', other records are identified by the year and quarter in which they were or are to be included.

Data is input into the SUCHAK database in one of three ways. Records of materials available on the premises are directly entered on the terminal. Information about materials in other libraries is obtained on worksheets, which are later input online. Still other records are created eternally on MICRO CDS/ISIS and dumped into SUCHAK using the ISOCONV processor. The distributed way in which the data is collected for SUCHAK makes it essential to edit the database at regular intervals and particularly before the running of an SDI service. While the daily editing identifies input errors, the periodic editing helps remove inconsistencies in terms and author names and checks duplicate records.

The READERS Database

The READERS database – which includes information about users – initially had a simple structure with only a few fields for name, address and number of copies needed as required by MINISIS. During the year we found ourselves consulting this database frequently and found it necessary to add other fields. It was very useful for administrative purposes to cross index individual users with their profiles, and a field to include their profile number was added. In many cases a brief explanation of the project, its purpose and duration for which the service was necessary. While this information was more appropriate to the interest profile of the user, it was placed in the READERS database as a profile was often shared by more than one user. In order to monitor the service, it was useful to link the user with an institution, but many users preferred to receive their notification at another address, such as their residence or place of work. Additional fields to provide the second address were found helpful.

Initially, none of the fields in the READERS database were inverted. During the implementation of the service we have found it useful to invert name, university/institution and city and state.

The PROFILES Database

Since the SDI service is to reach scholars all over the country, interest profiles are collected by a questionnaire which seeks information about title of project, objectives, methods, instruments and keywords. Local users are encouraged to visit the Centre to negotiate their profiles. We have also found it helpful to do a retrospective search using the terms supplied and sending a sample output for confirmation. Users' feedback and comments after receiving the alerting bulletin are also used for modification of profiles.

Profiles are being created for individual scholars, projects and institutions. The printing limit of institutions is set higher than that of individuals. The three subject areas are given different job numbers and different print formats to include only relevant fields belonging to the projected subsets of that subject.

The fields from SUCHAK most used for constructing the profiles are the subject descriptors. However, language, target audience code, orgin of document (Indian or foreign), geographic descriptors – all of which are inverted fields – have also been used. Occasionally we have also found it necessary to use free text searches in the statements.

All profiles include one search statement which limits the search to those records with the status "Current."

Certain topics and areas are found to repeatedly occur as interest profiles. Women and law, crimes against women, research in home science, women and health, child nutrition and consumerism are areas of frequent interest. Standard profiles in these subjects are generally shared by more than one user.

On the other hand, the interdisciplinary nature of the subject and their dependence on each other have made it necessary to construct more than one profile for a particular interest. Thus for example, a researcher working on parental attitude to childrearing practices or support systems used by working women may have their main profiles linked to home science or women's studies descriptors. However, since sociology can provide a theoretical base to understanding attitude formation and measurement of social networks and their roles, a second profile for the sociology subset in SUCHAK becomes necessary.

Since both the creation of the bibliographic database and the SDI processing is done internally, the close connection between the two has become evident to us. Any changes, additions and deletions in the terms used in SUCHAK, have repercussions on the PROFILES database. To locate the specific terms which need modification in the search statements a free text search has to be undertaken. This activity is likely to become more time consuming as the number of profiles processed increases. The inversion of the search statement field is thus being considered.

SDI PROCEDURES

To provide the SDI service, the query processor is used in the SDI mode and a set of profiles in one subject are matched against the SUCHAK database. Since all the profiles include a search statement for "Current" status, only the records of the relevant quarter are located. The outfile is then printed using the SDI processor. When SDI print for all the three subjects is complete the status of the records in SUCHAK is changed from "Current" to the the actual quarter e.g. 89/1. The status of the records of the next quarter are then changed to current. While most of these steps were undertaken in an interactive mode, they are now being combined into one job.

THE FUTURE

The SDI service was an unfamiliar one for both the Centre and for most of the users. In planning this new service, we opted to begin with a very simple but feasible model, hoping to refine it further in the light of experience and the feedback provided. The experience of the four bulletins has generated enthusiasm and confidence in us. We now propose to make a number of changes. As the SUCHAK database grows faster and more records are added in a month, the service will become bi-monthly. To improve on the relevance of the items notified, weighting will now be added. The use of 'any tables' to widen the searches are also planned. More effective print formats are being worked out. A changed version of the questionnaire used in the profile construction is also being experimented with.

Begun in a small way, the SDI service has provided us with an insight into our user needs in these fields. The fact that as many as half the users of the service are interested in extremely specific and local issues has assured us of the validity of the rationale of the service. As users experience the benefits of receiving notification of new materials the information about the Centre's new service spread, the number of new users registered goes up and the service gathers momentum.

Table 1 : The Subjects, their Orientation, Literature and Bibliographic services

	Sociology	**Women's Studies**	**Home Science**	**Library Information Science**	**Special Education**	**Gujarati**
Orientation	Disciplinary -Social Sciences	Mission oriented; interdisciplinary; drawing from subject areas in the social sciences and science and technology and to some extent from the Humanities.				Disciplinary - Humanities
Literature	Moderately coherent and large	Highly dispersed with a small core	Dispersed overall; but with coherent cores of varying sizes in the different specializations.	Moderately coherent and substantial	A small core with dispersed supportive materials.	Small and coherent
Regional Aspects of Literature used	Mainly local and culture based for facts and data and problems; international literature used chiefly for theories, concepts, methodology and comparisons.					
Indexing/Abstracting Services	International computerized abstracts available	Abstracts/ Indexes, Mainly in print form	Not available for the entire field; available for some specilizations; with not necessarily the required orientation.	International computerized abstracts available	Computerized abstracts available as part of a larger field.	Not available
Coverage of Indian Materials	Limited and late	Nil	Limited and late in some areas ; nil in others	Limited and late	Nil	N.A.
Academic Users	Small, coherent group at postgraduate and higher levels; large dispersed group in colleges	Small dispersed group of researchers	Small coherent group in special colleges; large dispersed group in general colleges	Small coherent group	Very small	Small
Other Users	Welfare Organizations	Welfare Organizations & Activists	Practitioners	Parctitioners	Practitioners	Journalists and Authors

The Chinese Data Entry with BATCHIN

Zhou Jiyu
Shanghai Population Information Centre

ABSTRACT

A file of non-ISO 2709 format can be imported into a MINISIS data base with BATCHIN processor. The procedure of Chinese data entry is different from that of English data entry.

Part one of this paper discusses the definition of the environment which must be made for the Chinese data entry with BATCHIN. Part two mentions two problems, which appear when BATCHIN is used for entering Chinese data and presents their solutions.

The BATCHIN processor of MINISIS document DBMS can be used to write records, which are in non-ISO 2709 format and reside on tape or disc, to a MINISIS data base. The function is very useful. We may import various formats of file, such as data from dBASE data bases used widely on personal computers, into a MINISIS data base using BATCHIN.

This paper discusses BATCHIN processes from the point of view of Chinese data entry.

I

First, the format of an input file which will be entered with BATCHIN must be in accordance with the format of the BATCHIN: at the beginning of the input file there are some control commands. The first physical record of each logical record must contain the record identifier "!REC-ID" only. Each field in a logical record is preceded by a field identifier "!tag!". etc. We may reform the input file with a programme written by in an advanced language if the record format does not match the BATCHIN format. The following is a sample of the BATCHIN format file.

```
!db=gc
!change strip=y,length=80
!rec-id
!a010!   530
!a020!B400200
!a030!30
!a040!上海市第三次人口普查分析资料选编
!a050!上海市人口普查办
!a080!1986
!a090!78
!a100!a
!rec-id
!a010!   531
!a020!B400242
!a030!06.14
!a040!社会和人口统计体系
!a050!联合国经济和社会事务部统计处
!a080!1985
!a090!490
!a100!b
```

Before running BATCHIN, we ought to set up the environment of BATCHIN, which contains two parts.

1. According to the actual cases of the user, set up three file equations.

The formal file of the BATCHIN input file is named BATCH and resides on tape. A file equation may be used to indicate that the input file's name differs from BATCH and the file is located on another device. For example, when the input file is named DATA and resides on disc, the following equation must be given:

```
:FILE BATCH = DATA;DEV = DISC
```

There are two possible output files from BATCHIN. One is an error listing, which is called ERRFL, another is a check listing, which is called OUTFL. The OUTFL is not produced unless the LIST = Y command is in the input file.

Both ERRFL and OUTFL are normally set to the system line printer. They can be redirected to other devices using the following equation:

```
:FILE ERRFL;DEV = < logical device number >
:FILE OUTFL;DEV = < logical device number >
```

2. Create a handler for the Chinese data entry with BATCHIN

Why must this step be done? Because the Chinese data must be processed with a handler before it is input into a MINISIS data base: the character strings of each character set are preceded with a special Character Set Identification Code (CSIC), and the eternal codes of the character set are translated into internal codes. After this process, the Chinese character strings will be recognize. Otherwise, we will get into trouble in string formatting, upshifting, and key extraction for Chinese characters. For example, although Chinese characters input as data can be viewed on the terminal in which a handler has been installed, the information cannot be retrieved from MINISIS. The number of hits of Chinese character free text searching is always zero. The Chinese character fast access search always results in "ILLEGAL TERM" error messages, even though it occurs in the B-TREE file.

Now, lets discuss how to set up a handler for BATCHIN.

There are some handler programs in the MINISIS segmented library (SL). Users ought to describe the device which will use Chinese characters in a specific file called TERMHDLR (Terminal Handler Table). However, the procedure of setting up a handler for BATCHIN is different than that for a terminal. We should use the pseudo logical device number and the Job Control Word (JCW). For example, we take 100 as our pseudo logical device number, and then do as follows:

1. Using the EDITOR subsystem, add the line "100HANDLER08000100" to the end of the TERMHDLR.PUB.MINISIS file.

2. Set up the JCW. The name of the JCW must be BATCH. The JCW value must be the pseudo logical device number used above. For instance, type the MPE command:

```
:SETJCW BATCH = 100
```

Thus the Chinese character handler has been created for BATCHIN. Before BATCHIN invokes the BATCH file, it will search the pseudo logical device number in the

TERMHDLR file with the JCW value as the key. Then the character strings in the BATCH file will be processed with through the handler which will:

1. Add CSIC.

2. Translate the external codes into internal codes.

All of this will be transparent to the user. Users can import Chinese characters with BATCHIN as easily as English.

After the environment for BATCHIN is set up, we can run BATCHIN. As soon as the data base name is entered, MINISIS will move the data from the BATCH file into the MINISIS data base.

II

Two problems which could appear during the importing of Chinese data with BATCHIN should be noted.

1. As we know, there are so many Chinese characters that two bytes must be used in the computer to identify one Chinese character. For this reason, we must use two bytes as a unit for processing Chinese characters. Sometimes there may be a "half of a Chinese character" in the BATCH file. For example, when we edit some records using the dBASE system, only one of the two bytes of a Chinese character are deleted. The byte left over cannot be recognized by the computer. The byte can be seen and deleted when using a white background in the dBASE system. Sometimes during the BATCHIN processing, the remaining byte with cause a dead halt. Users ought to review for and remove any remaining single bytes in dBASE before transferring the file for BATCHIN processing. Enlarging the length of the field will also prevent the dead halt problem but the unrecognizable character will be entered as data. Unrecognized characters can not be displayed on a terminal, but they can be printed on the line printer.

2. If we define the maximum length of the Chinese data field equal to its actual length we may get the "FIELD TOO LONG...FOLLOWING PART REJECTED" message during the BATCHIN procedure. Sometimes when the data is the actual length of the field, and contains characters such as punctuation and digits entered as non-Chinese characters, the message "FIELD TOO LONG" will occur.

The reason for this problem is that the handler inserts CSICs into the data to indicate the character set used. The CSIC occupies two bytes and is prefixed to the character string to identify the character set used by the string. The CSIC of the Chinese character set is octal 16 and 4. The CSIC of English is octal 16 and 0. If the CSIC is not present, the string is assumed to be in character set 0 (which is referred to as the default character set.) CSICs may be embedded in the middle of a string, indicating a change in character sets. Although a string in the default character set requires no CSIC at the beginning, a string which switches from a non-default character set to the default character set requires a CSIC indicating the change to character set 0. Thus it can be seen that we must consider the CSIC length of two bytes when we define the field maximum length. Furthermore, when there is punctuation or digits entered using a non-Chinese entry method, it will require more CSICs. For instance, if there is a punctuation mark entered using a non-Chinese entry method, two bytes must prefix the not only the punctuation, but the Chinese string following the punctuation. So when we define the maximum length of the MINISIS data base field, these prefixed CSICs must be fully considered so as not to lose data.

After MINISIS was installed on our computer, we imported the data from our old dBASE data base into our MINISIS data base. This makes it possible to use data in our dBASE data base to prevent data duplication, and to bring some effects of the data into full play. We used dBASE for checking duplication only. Data entered into MINISIS may share the strong retrieval abilities of MINISIS. If the data is further perfected, the data will provide users with an ideal information retrieval data base management system.

中文数据的 BATCHIN 输入

The Chinese Data Entry With BATCHIN

上海人口情报中心　　周基玉
Shanghai Population Information Centre
Zhou Jiyu

【内容提要】非 ISO2709 格式的数据文件可用 BATCHIN 模块输入到 MINISIS 库中．用 BATCHIN 传输中文数据的过程与传输英文数据不同．本文第一部分讨论了调用 BATCHIN 模块传输中文数据的环境设置，第二部分讨论中文数据传输过程中可能遇到的两个问题及处理方法．

【Abstract】 A file of non-ISO 2709 format can be imported into a MINISIS data base with BATCHIN processor. The procedure of Chinese data entry is different from that of English one.

Part one of this paper discussed the definition of environment which must be made for Chinese data entry with BATCHIN. Part two mentioned two problems ,which probably appeared when BATCHIN was being used for entering Chinese data, and their resolutions.

MINISIS 文献数据库管理系统的 BATCHIN 处理程序可将存于磁带或磁盘上的非 ISO2709 格式的数据输入到 MINISIS 数据库中，这个功能是非常实用的，借助它，我们就有可能将各种数据文件转人 MINISIS 数据库中．例如，微机上广泛使用的 dBASE 数据库，就可用 BATCHIN 模块将数据转人 MINISIS 数据库中．下面根据我们的实践谈谈使用 BATCHIN 模块的过程和传输中文数据时的特点．

一、建立符合 BATCHIN 模块输入要求的数据文件

首先，用 BATCHIN 模块批式输入的数据文件，应符合 BATCHIN 的格式要求：一开头是一些控制语句，记录之间用“! REC-ID”标识，字段以“! 字段名!”开头，等等．若数据文件不符合 BATCHIN 格式，可用高级语言改写后提供给 BATCHIN 模块．下面是一个 BATCHIN 格式数据文件简例:

```
!db = gc
!change strip = y,length = 80
!rec-id
!a010! ###530
!a020!B400200
!a030!30
!a040!上海市第三次人口普查分析资料选编
!a050!上海市人口普查办
!a080!1986
!a090!78
```

```
!a100!a
!rec–id
!a010! ###531
!a020!B400242
!a030!06.14
!a040!社会和人口统计体系
!a050!联合国经济和社会事务部统计处
!a080!1985
!a090!490
!a100!b
```

在运行 BATCHIN 模块将数据文件装入 MINISIS 库之前，要设置 BATCHIN 的环境，这包括以下两部分内容．

1、根据用户实际情况，建立三个文件方程．

BATCHIN 模块输入文件的形式文件名是“BATCH”，预置 BATCH 文件是在磁带上．若实际情况不是这样，例如数据文件名为 DATA，且 DATA 文件在磁盘上，就应建立如下文件方程予以引导：

```
:FILE BATCHIN = DATA; DEV = DISC
```

BATCHIN 模块输出错误信息文件的形式文件名是 ERRFL，预置输出到系统打印机上．在 BATCHIN 格式数据文件开头的控制语句中，若有“LIST = Y”，将输出查对信息文件，该文件的形式文件名是 OUTFL，预备输出到系统打印机上．对 ERRFL 和 OUTFL 这两个文件，在运行 BATCHIN 之前，可用下列 MPE 文件方程引导到用户所希望的设备（如终端）上：

```
:FILE ERRFL; DEV = 用户希望的逻辑设备号
:FILE OUTFL; DEV = 用户希望的逻辑设备号
```

2、为 BATCHIN 传输中文数据接“管理程序”（HANDLER）

接了管理程序后，从 BATCHIN 传输的数据就先经过管理程序处理：在各种语种字符串前加相应字符集标识码（CSIC）、将汉字串从输入代码转换为内部代码，然后才进入 MINISIS 数据库．经过这样处理后的汉字串能被 MINISIS 识别．若不经过管理程序处理，汉字串的倒排、索引就不能正确进行，例如进入库中的汉字串虽能在装了汉字管理程序的终端上显示，但无法检索：对非倒排字段的汉字检索，命中数总是零，尽管倒排文件中可以看到汉字对倒排字段的汉字检索，总是提示“非法关键字”．

在 MINISIS 的段库（Segment library）中原已放有管理程序，用户应在一个名为 TERMHDLR（终端管理程序描述表）的文件中对某台要作汉字输入输出的设备进行描述，此后在这台设备上通过的汉字就将经过管理程序的处理，可被 MINISIS 识别．

给 BATCHIN 模块接管理程序不同于给终端接管理程序，要用伪逻辑设备号和作业控制字（JCW）来实现．设伪逻辑设备号为 100，具体做法如下：

①在 EDITOR 子系统中修改名为“TERMHDLR.PUB.MINISIS”的文件，即在文件

末加一行"100HANDLER08000100"。

为避免修改 MINISIS 系统帐号下的文件，可将文件拷贝到当前帐号来修改，那就要改两个文件，做法如下：

: FCOPY FROM = TERMHDLR. PUB. MINISIS; TO = TERMHDLR. PUB; NEW

: FCOPY FROM = UDC.PUB.MINISIS; TO = UDC; NEW

然后在 EDTOR 子系统下给 TERMHDLR 文件加一行"100HANDLER08000100"，再删去 UDC 文件中"FILE TERMHDLR.PUB = TERMHDLR.PUB.MINISIS"这一行。

②建立作业控制字（JCW），JCW 的名称应是 BATCH，其值应是伪逻辑设备号，即在 MPE 状态下键人命令：

: SETJCW BATCH = 100

这样就给 BATCHIN 模块接上了汉字管理程序，在 BATCHIN 模块调 BATCH 文件时，先由 JCW 值检索 TERMHDLR 文件中的伪逻辑设备号，接上相应的管理程序，此后 BATCH 文件中的字符串就将经过管理程序处理：

1、加字符集指示码。

2、做输入代码与内部代码的转换。

这些对用户都是透明的，用户可以象用英文一样方便地使用汉字。

设置好 BATCHIN 的环境后，就可进人 MINISIS 的 BATCHIN 模块，然后给出 MINISIS 数据库的名称，系统就将 BATCH 文件的数据装人 MINISIS 库中。

二、运行 BATCHIN 模块传输中文数据的过程中的注意事项

1、由于汉字数目较多，在计算机内以两个字节表示一个汉字，所以处理汉字时应以两个字节为单位。在 BATCHIN 格式数据文件中，由于某种原因有时会存在"半个汉字"，例如在 dBASE 数据库中进行全屏幕编辑修改时，对一个汉字仅删除了一半，这样处于汉字状态，又只有一个字节，这剩下的一个字节就成为不可识别的字符，dBAES 对不可识别的字符不显示。在 dBASE 中用反显示方式可看到、删除这些不可识别的字符。MINISIS 的 BATCHIN 模块遇到数据文件中的不可识别字符，有时会造成死机，因此用户事先应删除数据文件中那些不可识别的"半个汉字"。用加长 MINISIS 数据库字段长度的方法亦可以避免死机，但不可识别字符也被装人数据库，终端对不可识别字段不显示，打印机输出时可看到不可识别字符。

2、对中文数据，若定义 MINISIS 数据库字段的最大长度为实际数据的长度，在 BATCHIN 运行时经常会得到"字段太长，部分数据丢失"的提示信息（信息号 M2272、M2273）。比如 dBASE 库中的一个字段定长长度为 40 个字节，我们定义 MINISIS 库相应字段的最大长度也为 40 个字节，运行 BATCHIN 后，当实际数据为 40 字节会出现"字段太长"的错误，另外，即使实际数据小于 40 字节，但其中有非汉字输入方式输入的标点、数字等，有时也会被提示"字段太长"。

这是由于汉字管理程序在处理时，为指明字符集类别，要在输入的数据中插入字符集

指示码CSIC。CSIC为两个字节，放在字符串之前，以指出其后的字符串来自哪一个字符集，亦即是哪一种文字。汉字的CSIC为%16%4(即八进制的16、4)，罗马字符集的CSIC为%16%0。若字段的一开始就是隐含字符集的字符串，则该字符串前没有CSIC，但如果一个隐含字符集的字符串紧跟在一个非隐含字符集的字符串之后，在隐含字符集号符串之前也应冠以CSIC。由此可见，定义字段长度时，对汉字串应考虑到这两个字节的CSIC长度，并且当汉字串中有非汉字输人方式输人的标点、数字等字符串而形成混合字符串时，就会有更多的字节用于CSIC。

如汉字串中出现一个非汉字方式输人的标点时，在标点前要有两个字节的CSIC，在标点后、汉字串前，又要有两个字节的CSIC。所以，在定义MINISIS数据库字段长度时，要充分考虑这一点，以免数据丢失。

我们在安装了MINISIS之后，将以前在dBASE下建的中文馆藏数据库的三千多条记录转到了MINISIIS数据库中。这个转换的完成，一方面使我们保护了以前的工作成果，避免重复的输人劳动，另一方面提高了数据库的功用，原来用dBASE建的数据库，只能做简单检索，而且速度又十分慢，基本上只能用于图书查重，现在数据转入MINISIS数据库后，能享用MINISIS的各种情报检索功能，若再对数据进行充实完善，就能作为一个理想的情报检索数据库提供给广大用户。

DEVELOPMENT AND MANAGEMENT OF AN AGRICULTURAL BIBLIOGRAPHIC INFORMATION SYSTEM

Franz Leemreize
Wageningen Agricultural University
The Netherlands

(1) INTRODUCTION

Although the Netherlands is a very small country in terms of surface area (about 4 million ha), and although it is densely populated (nearly 3.5 inhabitants per ha), agricultural income currently constitutes 10% of the net national income. Agricultural exports represent about a quarter of all national exports. The balance of trade for agricultural production is positive and growing. The success of agriculture in the Netherlands can be partly attributed to the agricultural library facilities and information services of the institutes of the Directorate for Agricultural Research and the Wageningen Agricultural University, which they provide through the expertise and collaboration of their staffs.

In this respect, the Netherlands is fortunate indeed, because together the nearly 80 libraries, among which the library of the Wageningen Agricultural University plays a central role, possess one of the most important collections of agricultural literature in the world.

However, it has been recognised for several years that the highly splintered organization of these collections restricts access to information, thereby reducing its potential effectiveness. Pressed, on the one hand, by government to rationalise services and, on the other, by users to improve library services to meet the challenge of the growing national and international information retrieval systems, we decided in 1977 to examine the requirements of an automated agricultural library and documentation system for Wageningen, where three-quarters of the agricultural libraries are located.

In that same year the libraries of the Wageningen Agricultural University and the Agricultural Research Institutes in the Netherlands set up a joint project to implement a shared automated library and documentation system. This system would contain information concerning the collections of nearly 8O, mainly small, libraries; provide about 2O libraries with on-line access to the information and functions of the system; generate library and documentation services from foreign databases; and provide interfaces to national and international information systems.

It was to be a first generation automated system for librarians in Wageningen. None of their activities had been automated and contact with other automated systems had been minimal. In 1978, it was decided to use the MINISIS relational database management system, developed by the International Development Research Centre (IDRC) in Ottawa, Canada, as a building-tool for this bibliographic system. MINISIS is very suited to bibliographic applications.

The resulting system is now called AGRALIN. This is an acronym for AGRicultural Automated Literature Information system Netherlands.

(2) AGRALIN

In Jan-Kopshuis (Jan Kopshouse) in Wageningen, and 55 other agricultural libraries in the Netherlands, AGRALIN is being used for library administration, storage and retrieval of scientific literature and documentation.

Much data are stored in the database, among these are 600,000 own titles and some 75,000 titles from the FAO. The system is characterized by the on-line accessibility of all the stored data from 130 terminals of participants and 90 terminals and PCs of scientific, commercial and industrial clients.

The system is still growing – there are still some 10 libraries to be connected to the system.

The first objective of the automated system is to integrate catalogues and documentation data of the libraries and documentation centres for agriculture in the Netherlands. The building of the terminal network is therefore closely linked with the establishment of an agricultural library network, on the one hand, and an agricultural documentation network, on the other hand. This is reflected in the organizational plan of AGRALIN.

The second objective of the system is to give the participating libraries all the automated functions necessary for efficient library management.

(2.1) SOFTWARE AND HARDWARE

The software used for AGRALIN comprises the MPE operating system, the MINISIS database management and retrieval system, and database structures, programs and procedures developed in-house the programs are written in SPL and FORTRAN.

Because the MINISIS software is only written for Hewlett Packard's HP3000 computer, the choice of our most important piece of hardware – the computer – was a combined one: MINISIS and the HP3000 computer. Within the range of HP3O00 computers, there are several models of different capacity that are fully software compatible. So, although the hardware choice was not an entirely free choice, it has turned out to be a good one. We are very impressed by the reliability of the hardware.

At this moment AGRALIN uses:

HP 3000 series 950 computer
  including 96 Mb of internal memory
3 lineprinters (400/600 1pm)
6 disc-drives (572 Mb each)
1 tape-unit (1600/6250 bpi)
130 terminals with local printers or barcode readers
several modems, multiplexers, linedrivers and PADs used for communication

(2.2) AUTOMATED FUNCTIONS

On 1 January 1980 the first functions of the system became operational. The last function implemented this year was the online catalogue of AGRALIN for the library public.

The main functions of AGRALIN are:

ORDER ADMINISTRATION

All orders for books and periodicals of the libraries of the Agricultural University are registered in the database. Every user of AGRALIN can see whether a publication is on order. The order forms and claims are printed by AGRALIN; all financial data are transferred to financial-administrative systems.

CATALOGUING

The database contains information on titles and data about the location of the publications. This function is essential to the information services since it makes it possible to compile an integral catalogue. Shared cataloguing and shared use of catalogue information decreases the input time and increases the efficiency of bibliographic research.

STORAGE OF DOCUMENTATION DATA

Documentation records of publications selected from agricultural journals of interest can be stored in the database, to be later retrieved and printed on a laser printer to produce camera-ready copy for 25 different current-awareness bulletins.

CLASSIFICATION

Universal Decimal Classification (UDC) is used to classify documents in the system. A special, centrally controlled UDC database is used to manage the classification of all documents.

SEARCHING FACILITIES

It is possible to search for information stored in the database in four ways:

on-line, using the standard query function of MINISIS, which is a powerful information-retrieval tool.

on-line, using a menu-driven query function.

The main difference between this and the standard on-line query function is that end users of a library and other inexperienced users do not need to know any command language.

in the catalogue microfiches supplied by AGRALIN; these are the personal author names and the title catalogues for books; the systematic UDC (Universal Decimal Classification) catalogue for books and documentation; and the title catalogue for periodicals.

in several of the printed listings provided for library staff.

LOANS ADMINISTRATION (LOANS)

The loans administration manages the lending out, reservation, and return of items, and at the same time can provide information directly from the catalogue. Reminder cards are printed automatically by the system. The "loans" is based on the use of barcodes for borrower and item identification.

SERIALS ADMINISTRATION (CARDEX)

The serials administration registers the incoming numbers of a running serial. In cases of discontinuity it automatically prints and sends claims. CARDEX consist of a registration procedure, a claims procedure and some other administrative procedures. Internal circulation can be used as a part of CARDEX.

SDI (Selective Dissemination of Information)

The SDI system has been operational since 1983. With this system, tapes from other systems with information about new titles are regularly searched at the request of users. Search profiles determine what information is to be selected for the user. The SDI is implemented for profiles to run against information coming from different external databases. These are CAB, AGRICOLA, FSTA and AGRIS. Of course, SDI also runs on AGRALIN's own bibliographic database.

HOST FUNCTION

For persons or institutions that do not participate in AGRALIN, we have a special function to give them access to bibliographic information. These users are charged for connect time. They can access us using dial-up lines, hired lines or the national packet-switching network.

(3) MANAGEMENT OF AGRALIN

There are two phases to be considered:

> project management during the development and implementation of the system;
>
> management of an operational system.

(3.1) PROJECT MANAGEMENT

The success of every automation project can be represented by a progress curve. On the y-axis of the graph you can see the amount of progress as function of the time. The progress planned is represented by the upper line. The actual progress is represented by the lower line. The planned-progress curve depicts a smooth transition from the software development to implementation. In reality, however, at the implementation point project workers are confronted with demotivating events, such as:

1. The users are dissatisfied, because they get something different to what they expected.
2. Users have become 'automation minded' and they ask for other or better system features.
3. The system software is not flexible enough, or separately developed sub-systems cannot be easily integrated.
4. System characteristics that are difficult to predict, e.g. run time and response time, turn out to be worse than expected.
5. Simple – yet fundamental – deficiencies in the systems analysis become apparent.

The period between the implementation point and 'window point' is the demotivation time. During this time the project team is patching up wounds by providing the missing links in the system and cleaning its face with minor local software changes. At 'window point' there are two paths to choose from: either throw the system out the window (hence the term 'window point'); or evaluate the entire system and rebuild it using the components that are good. The last choice will combine the efforts of users and builders to produce a system that meets the real requirements made of it.

How can we shorten the demotivation period? Fortunately there are some ways to do this.

1. Before starting on the real development do some prototyping,

to expose the users to the system and to let them see how it will effect them, e.g. inputting procedures and the output products.

2. Let the users test their part of the system at a very early stage and get them write their own user documentation, so they can discover deficiencies in the procedures.

3. Realise that any change in the system concepts can have more consequences for a user than a systems analyst can imagine.

4. Project workers should be aware of the inevitability of the demotivation period and accept it as a part of their professional experience.

(3.2) MANAGEMENT OF AN OPERATIONAL SYSTEM

When a system is operational the management of the system has two main tasks to fulfil. First, there is the responsibility of the automation department for the technical "health" of the hardware and software of the system.

This can be met by having a system manager responsible for the hardware, data communication and output (products) of the running system and a software manager responsible for the software and the technical database management. Because AGRALIN is "on air" 24 hours a day, the system manager is assisted by 2 operators and the software manager by 1 programmer.

Second, there is the responsibility for the user procedures and the contents of the databases. These tasks can be best executed by users that have the most affinity with a subsystem or function within the system.

In the case of AGRALIN, the Automation Department of Jan Kopshouse is responsible for the technical aspects of AGRALIN, whereas the heads of the other departments of Jan Kopshouse are each responsible for their own part of the system. For example, the head of the Cataloguing Department is responsible for the cataloguing procedures and the contents of the Catalogue database, whereas the head of the department Public Services is responsible for the "loans" module, including the contents of the Item database and a Borrowers database, and the on-line catalogue facility.

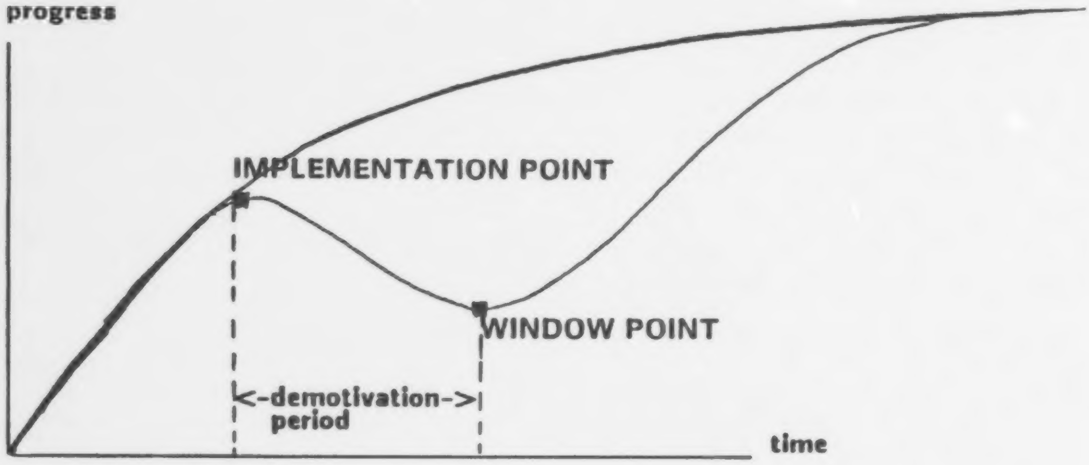
(4) PLANS FOR THE FUTURE

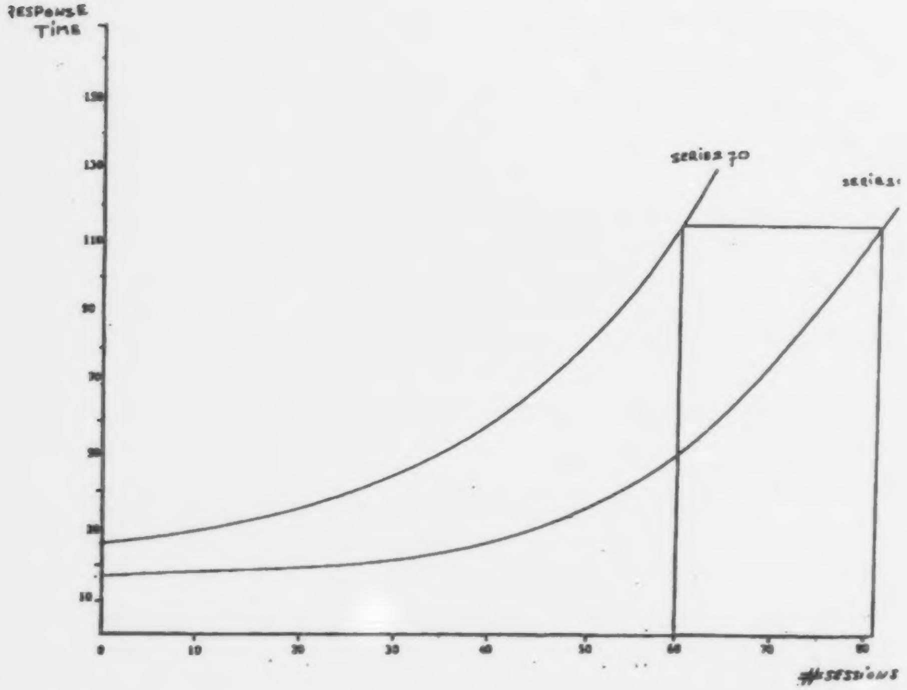
In the near future, three main activities are planned within the scope of AGRALIN. We will upgrade the HP 3000 to a series 955 computer, to add more power and give more users access to the system.

Furthermore, we intend to develop programs for storing the entire bibliographic database on CD-ROM, to make it possible to retrieve the data with a personal computer. CD-R0Ms would also supersede microfiche output.

Other facilities, for example on-line document ordering, will also be provided in the future.

As you can see, although a system appears complete at a particular point in time it will continue to change and evolve – otherwise it will die.


progress
IMPLEMENTATION POINT
WINDOW POINT
<-demotivation->
period
time



			JULY 1989		OCT.1989		? 1990/1991
SYSTEM	HP 3000 series 70	———>	HP 3000 series 950	———>	HP 3000 series 955	———>	
MINISIS VERSION	VERSION G		—				VERSION H
OPERATING SYSTEM	MPE-V		MPE-XL Comp.mode		MPE-XL Comp.mode		MPE-XL Native mode
EXTERN MEMORY	5 X 404 Mb		6 X 571 MB				
INTERN MEMORY	6 Mb		96 Mb				
MAX #SESS	60		80		120		180
RESPONSE TIME	POOR		BETTER -MEAN 50%-		50% BETTER		50% BETTER
THROUGH-PUT	1		1,3		2,0		3,0

INCREASE INPUT SPEED OF CHINESE CHARACTER BY USE OF BATCHIN

Wu XiaoHong
Instrumentation General Technology and Economics Institute
Ministry of Machinery and Electronics Industry

[abstract] BATCHIN processing program can be used to load in non-ISO format into a MINISIS data base. Data may come from an MPE EDITOR file, a non-MINISIS data base or a MINISIS data base and reside on tape or disc. Using BATCHIN can solve the problem, that is the speed or Chinese character data input. If the length of physical record is extended, how does it process. There are some questions in using BATCHIN.

I SUMMARY

Domestic and foreign instrumentation products data base is set up on the MINISIS data base management system. It is a large-scale Chinese character data base, having 12 million characters. The HP-VECTRO terminal we use has a few ways to input, and it's not widely used now, the way of five typeface input has it's characteristic high speed input, easy-learning and use can't be used at this terminal. For speeding up the input velocity. We entered data into dBASE III, then use the function of BATCHIN to input data into Minisis data base without adding any other equipment.

II BATCHIN Processing Program

BATCHIN expects the input file to be built on the basis of set standards. The length of the data in bytes in the physical record must not exceed 256 bytes. The logical input record may extend across several physical input records. REC-ID is used to delimit logical records in the input file, therefore it must be present for each logical record. The [ISN-number] is specified only when the data base is MASTER and auto-numbering is off. If the input file comes from a microcomputer, it must be processed through a terminal handler program. This reformats the informatoin so its internal format accords with the internal format of MINISIS.

III BATCHIN Input File Format

1. Special Identifier

!DB = < name >

< name > is the name of the MINISIS data base to which the input file will be written. If BATCHIN is run interactively, < name > can be omitted, or overridden by a command at the terminal.

!change command(s)

hange command redefines default characteristics of the batch input record. If change is not entered, the default values will be supplied. When running interactively, the hange command can be overwritten from the terminal.

The change commands are:

STARTER=

The default values is !. This is any character other than a comma or a blank space, used to indicate that the field-id immediately follows.

ENDER=

The default value is ! .This is any character other than a comma or a blank space, used to indicate the end of the field-id.

STARTER and ENDER characters may be other than for the field-id only, not for the !DB, !REC-ID and !change commands; they are always preceded by !.

Strip=Y/N

The default value is no. If it is changed to yes, the trailing blanks are stripped from every physical record.

Length=

The default length is the physical record length of the input file. Length and DISP combined must not exceed 256 bytes.

DISP=

The default value is 0. This is the displacement relative to 0 of the beginning of the data in the physical record.

EXIT=

The default value is None. A user-defined exit may be written to convert special physical records to the standard logical record format.

LIST=Y/N

The default value is No. IF the user specifies list=Yes, a full listing will be produced of each record created from the batch input file. If the user specifies list=No, no check listing will be produced.

!Rec-id [ISN-number]

Rec-id is used to delimit logical records in the input file; therefore, it must be present for each logical record. The [ISN-number] is specified only when the data base is MASTER and autonumbering is off—this number becomes the ISN of the MINISIS record.

2. Elementary Field

The field may extend across several physical records. The field identifer must be present for each field, it occupies the first place in the field.

Ex1:

```
!A100! Canada Today !A200! ABCD
```

Ex2: (with Strip=Yes)

```
!A100! Canada Today
!A200! ABCD
```

3. Repeatable Elementary Field

Occurrences of repeatable field are entered individually.

Ex: (with STRIP Yes)

```
!B010! MARY
!B010! Tom
!B010! Henlen
```

4. Subfielded Field

Subfielded field is entered as above, However, the first subfield of each subfielded occurrence must be preceded by the field start signal immediately followed by the subfield field-id.

Ex: (with STRIP = Yes)

```
!B200
!B201! ABCD
!B202! OTTAWA
```

5. Repeatable Subfielded Field

Occurrences of repeatable subfielded fields are entered individually.

Ex: (with STRIP Yes)

```
!B200
!B201! ABCD
!B202! ¥20.00
!B200
!B201! Canada Today
!B202! ¥30.00
```

IV Sample Input Files of BATCHIN

```
!DB = CPK
!Change strip = yes, length  = 80
!Rec-id
!A100! Computer development
!A200! China
!Rec-id
!A100! Canada Today
!A200! Canada
```

V Running BATCHIN in Batch Mode

1. Input Files

We use the data entry method of Five Typeface to enter data into a dBASE data base, then, periodically, we transfer data from dBASE III to MINISIS.

There is no record identifier and field id in the outfile of dbase III. The data base identifier and field id are written into the outfile by a processing program. When the input file format is the same as standard input file of BATCHIN, it is changed from external code to internal code by running the Xcode program and transferring the file into the central computer system.

The input file may be an unnumbered MPE EDITOR file, a non-MINISIS data base, a MINISIS data base which has been sent to disc or tape via the PRINT processor, or any file which can be massaged into the format that BATCHIN expects. The input file consists of records containing data to be entered into the data base, and records containing commands to the BATCHIN processor.

Each command must start on a new physical record. The data portion of the input file consists of a record identifier together with the field identifer and data to be associated with that record. Commands must start in column 0, but field identifiers and data may begin at a position specified by the DISP command.

2. Enter Data Using BATCHIN

The input file name is random. BATCHIN expects the input file to be named BATCH and reside on tape. A file equation may be used to indicate that input file is coming from some other device or the input file name is not BATCH.

If a input file name coming from microcomputer is YDCPK and it resides on disc and the length of physical records are not greater than 256 bytes, then the way of running BATCHIN to enter data is this:

```
:FILE BATCH = YDCPK;DEV = DISC
:SETJCW BATCH 999
:RUN BATCHIN.UTILITY.MINISIS;LIB = P
```

or

```
:FILE BATCH = YDCPK;DEV = DISC
:SETJCW BATCH 999
:MINISIS
check – 13
```

3. The Action and Usage of Terminal Handler Program.

One of the function of TERMHDLR is to transform the character string coding plan of MINISIS internal format into the display character coding plan of external device, or input character string coming from terminal into internal format of character string of MINISIS.

The standard character string of Minisis is made up of two parts: CSIC, character code string. If there is no CSIC in the front of characters string, the characters string belongs to default character set. CSIC may be located among the characters string and indicates the characters string changes. When the first character of the character string belongs to default character set, no CSIC is inserted. If the characters string changes from non-default to default, the front of default characters string must be identified by a CSIC marker.

The input file coming from microcomputer must be processed through TERMHDLR to insert the CSIC markers and transform its format into MINISIS internal format before BATCHIN processor is used to input data.

Because the logical device number of a file using terminal is anticipated, a record must be set up in the TERMHDLR table (Ex: 999 Handler08000100), with a fake logical device number (Ex: 999 is fake logical device number). A Job Control Word (JCW) is built at the same time. The JCW value equals the fake logical device number. JCW's name is the same as that of the in-out file. MINISIS will always look first for a Job Control Word established under the name of a terminal or line printer. If a JCW exists, then MINISIS will use its value as the logical unit number and will search the Terminal Handler Table for the logical number. HANDLERnn is called. If a JCW does not exist, then the logical unit number assigned by MPE will be used to search this table.

4. Change Field Length of MINISIS Data Base

For Chinese data base, the character string attribution is changeable. Owing to that a CSIC occupies 2 bytes, the field length must be enlarged to ensure that data isn't lost. Thus it can be seen that the field length of MINISIS is larger than one of dbase III.

5. Enter Data in Batch Mode

If BATCHIN is run interactively, the terminal is placed in a wait state.

In order to shorten the wait time, BATCHIN is run in batch mode. So a job is started, you may not wait to the end. When running BATCHIN in batch mode, the input file must contain the correct <name> and a CHANGE change statement if values other than the default are to be supplied.

A sample batch job might look like:

```
!JOB BATHW.MGR.CPK
!FILE MESS00.PUB=MESS00.PUB.MINISIS
!FILE ERR00.PUB=ERR00.PUB.MINISIS
!FILE SYN00.PUB=SYN00.PUB.MINISIS
!FILE TERMHDLR.PUB=TERMHDLR.PUB.MINISIS
!FILE BATCH=YDCPK;DEV=DISC
!SETJCW BATCH 999
!RUN BATCHIN.UTILITY.MINISIS;LIB=P
CPK
EXIT
!EOJ
```

BATCHIN will validate fields against authority files, but the data base must be in the character mode.

VI Surpass Length Record

For the input file coming from microcomputer, the length physical record may reach 2000 bytes. This record doesn't enter immediately. Before data is entered, the physical record length must be revised according to the demands of MINISIS.

Because a logical field may extend across several physical record, changing physical record length is easy. There are two ways to revise length. One is running a program, the other is using MPE command.

1. Running Program

If the physical record length isn't considered, processing the excess record length is easy, every filed contents are entered in a row.

Ex:

The length of A100 in MINISIS data base is 800 bytes, then A100 is divided into four parts:

A11, A12, A13, A14 in dbase III data base. The length of A11 is 200 bytes. The same is true for A12, A13, and A14. The contents of A11 come first, A12 is second, A13 is third, and A14 is last, when they are written in a row, data form is:

```
!A100! first second third last
```

The physical length of data is 800 bytes at the moment.

If the length of data doesn't exceed 256 bytes, then their fields are written individually, the form is:

```
!A100! first
second
third
last
```

The physical length of data is 200 bytes.

2. Using MPE Command

If the physical record of input file residing on tape or disc exceeds 256 bytes, the length may be changed by MPE command. It would be best if new length is common divisor of original length.

Ex:

```
:FILE OLDFILE;DEV = DISC;REC = -800,1,F,ASCII
:FILE NEWFILE;DEV = DISC;REC = -80,1,F,ASCII;ACC = INOUT
```

Using Fcopy Command again.

VII The Problem

1. String not Found in Text

When the character string is modified, the message "STRING NOT FOUND IN TEXT" may occur even when the character string may be seen.

2. Non-Regular Character

Something unexpected happens in code conversion, non-regular characters are copied but not converted. It might lead to a error in the character string in MINISIS. When the

MINISIS data base is in character mode, it is impossible to change the incorrect character string.

Conclusion

Because we use the way of Five Typeface Input and full screen editing, the input-velocity is 2 times faster than the previously. The microcomputer is used effectively and the minicomputer time is shortened, their advantages are fully effected.

利用BATCHIN功能提高汉字输入速度

吴晓红

[提要]Minisis中的BATCHIN功能是以交互或批式运行的方法,将早已存在于磁带或磁盘上的非Minisis格式的记录送入数据库.本文介绍用BATCHIN功能解决大量汉字输入速度问题,物理记录超长数据的处理方法,以及在使用BATCHIN中所遇到的问题.

一、概述

在MINISIS数据库管理系统上建立的国内外仪器仪表产品数据库是一个大型的中文数据库,大约1200万字符.由于我们输入用的终端是HP-Vector,输入方法少,使用不普及,而且目前输入速度快、易学、易掌握的五笔字型输入法在这种终端上又无法使用。为了提高输入速度,我们在不增加设备的情况下,采取了将数据输入到DBASE Ⅲ 中,再利用BATCHIN功能把数据批式输入到MINISIS数据库中。

二、BATCHIN处理程序

BATCHIN的输入文件必须按一定标准创立,物理记录可小于或等于256个字节.一个逻辑记录可以存放在几个连续的物理记录中.物理记录用!REC-ID分割,如果MINISIS数据库不是自动编号的,但它是一个主数据库,那么BATCHIN输入文件必须提供ISN.如果BATCHIN的输入文件是从微机来的,要先经终端管理程序处理,使其内部格式符合MINISIS要求。

三、BATCHIN输入文件特点

1.特定符号

!DB=MINISIS数据库名

如果运行BATCHIN程序,必须提供数据库名,你也可以在终端上提供或超越此名.

!CHANGE 修改命令

可以任选,修改下述参数的隐含值:

STARTER=

隐含值是!,指明字段标识符开始

ENDER=

隐含值是!,指明字段标识符结尾

STRIT=Y/N

隐含值是NO,如果是YES,在物理记录中数据之后的空格将被剥去.

LENGTH=

物理记录长度的隐含值,你可以改变,但不能大于256个字节.

DISP=

隐含值是零

EXIT=

隐含值是NONE

LIST=Y/N

隐含值是NO,如果LIST=NO,对于OUTFL来说什么也不处理.

!REC-ID

必须提供,用来分割一组作为一个逻辑记录的物理记录.

2.字段

一个逻辑字段可以延伸到几个物理记录,被写到MINISIS数据库的数据字段标识符居先.

例如: !A100!Canada Today!A200!ABCD

或(使用STRIP=YES):

!A100!Canada Today

!A200!ABCD

3.可重复字段

分开地写重复字段的每一个出现.

例如: !B010!MARY!B010!TOM!B010!HENLEN

或(使用STRIP=YES)

!B010!MARY

!B010!TOM

!B010!HENLEN

4.带有子字段的字段

首先送入0级字段标识,紧跟第一个子字段,数据,第二个子字段,数据……

例如:(使用STRIP=YES)

```
!B200
!B201!ABCD
!B202!OTTAWA
```

5.可重复的含有子字段的字段

分开地送入每一个出现

例如:(使用STRIP=YES)

```
!B200
!B201!Computer development
!B202!￥20.00
!B200
!B201! Canada Today
!B202!￥30.00
```

四、BATCHIN输入文件样例

```
!DB=CPK
!CHANGE STRIP=YES,LENGTH=80
!REC-ID
!A100!电子显微镜
!A200!D2-A型
!REC-ID
!A100!激光器
!A200!JG-W型
```

五、使用BATCHIN处理程序输入数据

1.输入数据文件的来源及要求

我们使用五笔字型输入方法,将数据输入到DBASE Ⅲ中,定期传送到MINISIS数据库中.

在DBASE Ⅲ中输出文件中是没有头标信息及字段标识符的,我们通过一个程序来处理DBASE Ⅲ输出的TXT文件,加入头标信息及数据相对应的字段标识符,使其程序处理后的数据文件符合BATCHIN的要求.再经过代码转换,将文件送到主机中.

输入文件可以是没有编号的MPE EDITOR文件,一个非MINISIS数据库文件,或一个能变

成BATCHIN格式的其它文件.这个输入文件由二部分组成,其一是包含数据的记录,其二是送给BATCHIN处理程序的包含命令的记录.每个命令必须是一个单独的物理记录,输入文件的数据部分由记录标识、与数据内容连在一起的字段标识组成,命令的开始列必须是0列,而字段标识和数据的位置可以通过DISP命令指定.

2.使用BATCHIN输入数据

从微机送到主机上的输入文件名是随机的,BATCHIN使用的标准输入文件名为BATCH,并常驻在带上;如果你使用的输入文件名与BATCHIN标准输入文件名不同,则必须提供一个文件方程式来说明.

假设从微机上来的输入文件名为YDCPK,驻在磁盘上,且物理记录长度小于或等于256个字节,那么运行BATCHIN装入数据过程如下:

```
:FILE BATCH=YDCPK;DEV=DISC
:SETJCW BATCH=999
:RUN BATCHIN.UTILITY.MINISIS;LIB=P
```

或:

```
:FILE BATCH=YDCPK;DEV=DISC
:SETJCW BATCH=999
:MINISIS
```

选择BATCHIN功能模块

3.终端管理程序在BATCHIN处理程序中的作用及使用方法

终端管理程序功能之一是将MINISIS内部格式字符串转换成在外部设备上显示字符串需要的编码方案,并将来自终端的输入字符串转换成MINISIS内部字符串格式.

标准MINISIS内部格式的字符串是由CSIC及字符代码串组成,如不存在字符集标识符CSIC,那么字符串是隐含字符集,CSIC可嵌入字符串中央,指出字符串的改变.在以隐含字符串开头的字符串之前不加入CSIC,但如果从非隐含字符串改变到隐含字符串,在隐含字符串之前也插入CSIC.

对于来自微机的数据文件,必须经过终端处理程序,插入CSIC,将其变成MINISIS的内部格式,再装入数据库.

由于输入输出文件所在终端的逻辑设备号是不能预知,因此必须先在终端管理程序表

中建立一个记录(如:999HANDLER08000100),它具有伪逻辑设备号(如:999就是伪逻辑设备号),同时建立一个作业控制字(JCW),使其值等于伪逻辑设备号.JCW的名称应与输入输出文件名相同.MINISIS检索TERMHELR文件时首先看一看JCW存在否.如有,则使用其值作为逻辑设备号检索TERMHDLR,找出对应的逻辑设备号.调用终端管理程序.

4.改变MINISIS数据库的字段长度

对于中文数据库,字符串属性变化可能很频繁,因此,你所定义的字段长度就应考虑适当放大,以保证数据不会因加了CSIC超长而被截断.因为一个CSIC占用二个字节.

也就是要MINISIS定义的字段长度一定大于DBASE Ⅲ所定义的字段长度.

5.批式输入数据

如果交互式运行BATCHIN则要占用终端很长时间,为了不使人等待时间长,可以批式方式运行BATCHIN,你只要启动一个作业,就可以做另外的事情而不必等它结束.当批式运行时,在输入文件中必须提供正确的数据库名与修改命令.

作业样例:

```
!JOB  BATHW,MGR.CPK
!FILE  MESS00.PUB=MESS00.PUB.MINISIS
!FILE  ERR00.PUB=ERR00.PUB.MINISIS
!FILE  SYN00.PUB=SYN00.PUB.MINISIS
!FILE  TERMHDLR.PUB=TERMHDLR.PUB.MINISIS
!FILE  BATCH=YDCPK;DEV=DISC
!SETJCW  BATCH=999
!RUN  BATCHIN.UTILITY.MINISIS;LIB=P
CPK
EXIT
!EOJ
```

BATCHIN可以执行字段合法验证功能,但如果数据库处于全屏幕方式,则无法执行此功能.如果有合法验证字段,在使用BATCHIN之前应先使数据库处于字符方式.

六、物理记录超长的数据文件的处理

从微机上来的输入文件,它的物理记录长度可达2000字符长,这显然不能直接使用

BATCHIN.如果要输入这样长的数据,则必须改变它的物理长度,使其符合BATCHIN的要求.

改变方式有二种,其一是在微机上通过程序来完成,其二是在主机上通过MPE命令完成.由于一个逻辑字段可以延伸到几个连续的物理记录中,所以改变物理记录长度则变得很简单了.

1.程序法

设一个字段长度为800字符,那么对应DBASE Ⅲ 中至少要4个域来满足,做为一个完整的字段.如果不考虑物理长度的限制,程序处理则很简单,只要连续写入既可.

例如:A100为800字符对应A11为200字节,A12,A13,A14也分别为200字节. A11内容为FIRST,A12为SECOND,A13为THIRD,A14为END,那么写入数据文件后则变成:

```
!A100!FIRST  SECOND  THIRD  END.
```

数据的物理记录长度为800字节.

如果使物理记录长度不超过256个字节,那么程序在写入数据时必须分开写.形式如下:

```
!A100!FIRST
SECOND
THIRD
END
```

这样数据文件的物理长度就是200字节长了.

2.MPE命令法

如果一个在主机中的数据文件物理记录长度超过256个字节,则可以通过MPE的命令将其改变成小于或等于256个字节.改变后的长度最好为原长度的因子数,以确保数据不会丢失.

例如:

```
:File  oldfile;dev=disc;fec=-800,1,f,ascii
:File  newfile;dev=disc;rec=-80,1,f,ascii;acc=inout
```

再使用Fcopy命令,就可完成这一个要求.

七、使用BATCHIN处理程序所遇到的问题

1.字符的不识别问题

在DBASE Ⅲ 中输入的字符在送入到MINISIS数据库中后修改时偶而出现没发现字符串

的问题.

2.非正规字符的输入

从DBASEⅢ送到MINISIS中的数据,常会遇到一些非正规字符仅考贝而未转换的问题.在MINISIS数据库中有时会引起大亮块或一串错误字符.在字符输入方式下,无法修改错误字符串.

八、结束语

由于运用DBASEⅢ的全屏编辑、五笔字型输入方法、输入速度提高一倍多,也充分利用了现有微机的作用,减少因输入而占用大量的主机机时,使它们各自的优点得以充分发挥.

Minisis in King Faisal Center

By Ibrahim Salem

INTRODUCTION

The King Faisal Center was established by the King Faisal Foundation in 1983 for the service of Islam and muslims everywhere. The objectives of the Center derive from the general aim of the King Faisal Foundation i.e the service of Islam. The Center aims at promoting scientific research and cultural activities beneficial to Islam and Islamic civilisation. It is mainly concerned with expounding the contribution of Islamic civilisation to human endeavours in the various fields of life throughout history, highlighting the main characteristics which distinguish Islamic civilisation from others, encouraging the development at scientific research in the fields at Islamic civilisation, demonstrating the humane qualities and the leadership of the late king Faisal, contributing to the preparation at competent scholars, promoting Islamic heritage, organising lectures on the issues vital to muslims, collecting manuscripts, and establishing a specialised database on computer to serve muslims' needs.

The Center owned two HP3000/3 and HP3000/58, we began using MINISIS about 6 years ago, we developed our own system in cooperation with ALDOC, we had to rewrite the message file and the syntax file in ARABIC and to use the function keys so the user cab use MINISIS in full Arabic. We are using MINISIS in:

a) Center libraries
b) Aduvtual library
c) Manuscript library
d) databases

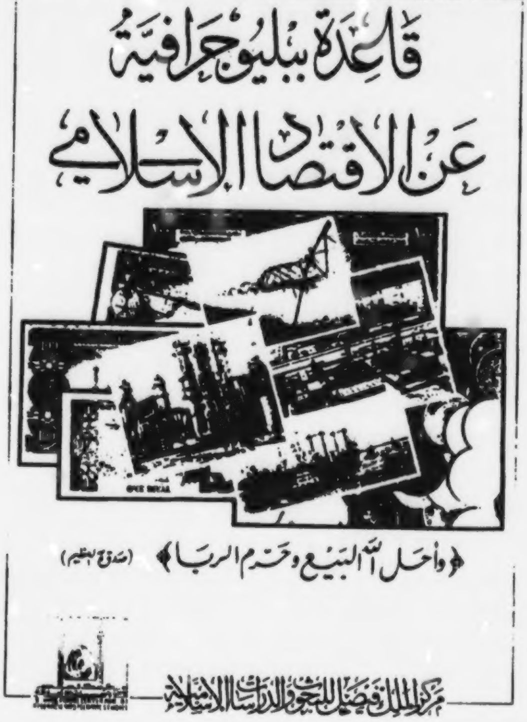
and some special perposes.

The library holding is more than 90,000 titles of books in ARABIC and other titles for periodicals and manuscripts. The center has created 12 databases in special fields like King Faisal, working women, and Islamic education and we developed new databases in university theses and muslims manorty.

The center has it own press so we interface the HP3000 with Crtronxs to print our index and reports in printing format. We printed three issues of manuscripts and we just printed Islamic theses in two volumes in ALPHPITIC orders and hierarchy. We are in the process of printing an index for all the University theses given by 7 Universites and 5 colleges in SAUDI ARABIA.

We have users inside the center in RIYADH and users outside Riyadh like Mecca, Jedah where we installed Tevmenal in Public libraries and universities, we gave advice and help to MINISIS users in SAUDIA ARABIA.

We are working now on interfacing the HP3000 using MINISIS in ARABIC with MacIntosh computers to use the desktop publishing tools available on the MAC II.

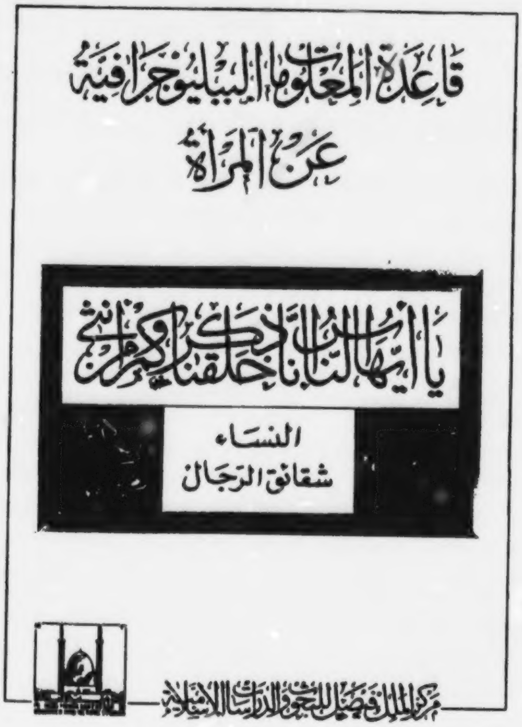


قاعدة ببليوجرافية

عن الاقتصاد الإسلامي

﴿وأحل الله البيع وحرم الربا﴾ (صدق الله العظيم)

مركز الملك فيصل للبحوث والدراسات الإسلامية



قاعدة المعلومات الببليوجرافية

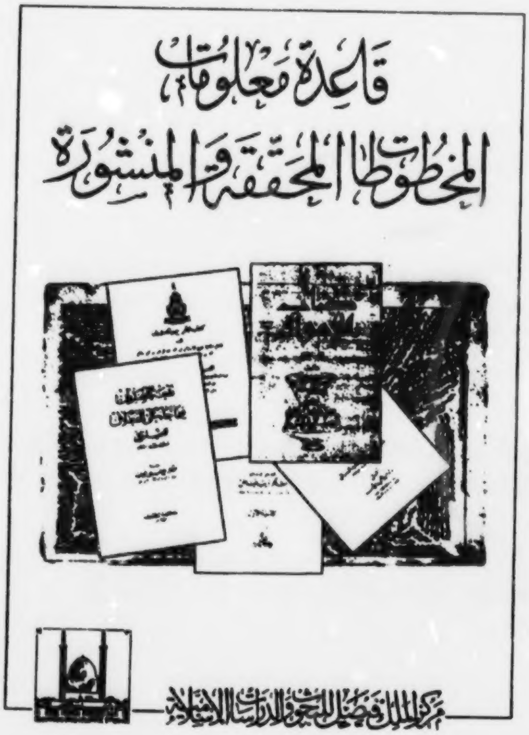
عن المرأة

يا أيها الناس إنا خلقناكم من ذكر وأنثى

النساء

شقائق الرجال

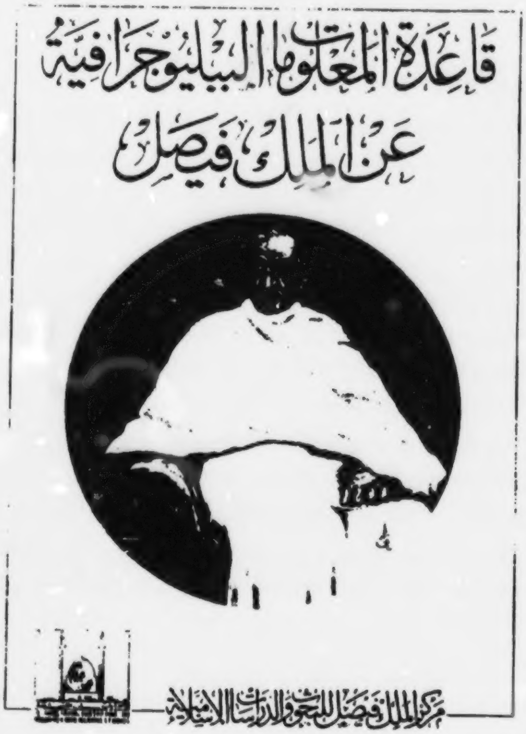
مركز الملك فيصل للبحوث والدراسات الإسلامية



قاعدة معلومات

المخطوطات المحققة والمنشورة

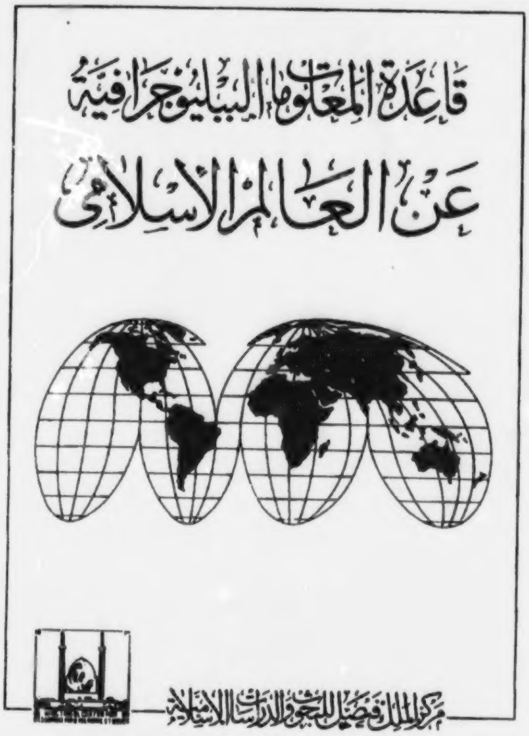
مركز الملك فيصل للبحوث والدراسات الإسلامية



قاعدة المعلومات الببليوجرافية

عن الملك فيصل

مركز الملك فيصل للبحوث والدراسات الإسلامية



قاعدة المعلومات الببليوجرافية

عن العالم الإسلامي

مركز الملك فيصل للبحوث والدراسات الإسلامية

قاعدة المعلومات المتخصصين

في الدراسات العربية والإسلامية

وما أوتيتم من العلم إلا قليلا

مركز الملك فيصل للبحوث والدراسات الإسلامية

قاعدة المعلومات الببليوجرافية

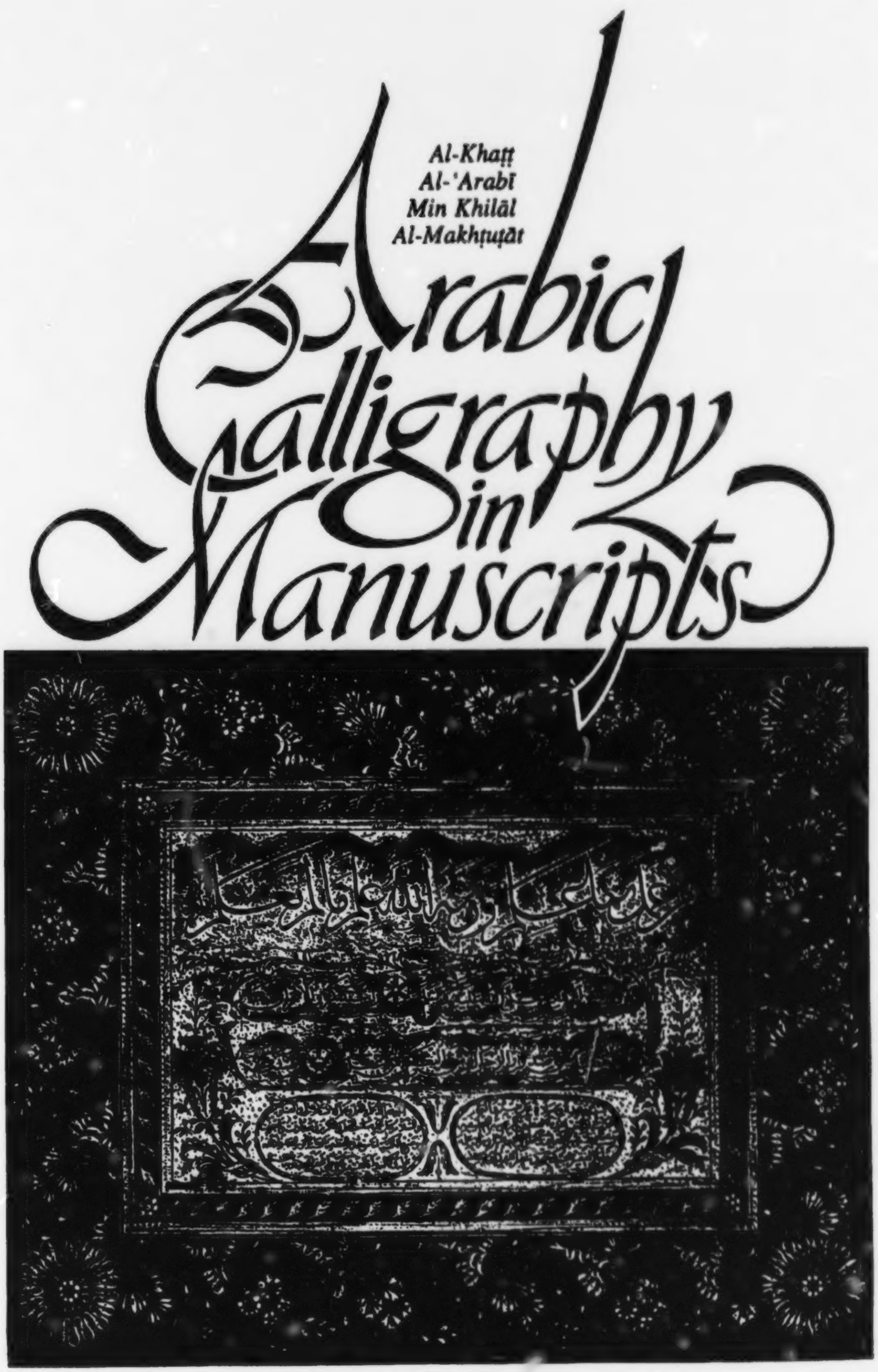
عن الإعلام الإسلامي

مركز الملك فيصل للبحوث والدراسات الإسلامية

قاعدة معلومات الدول

والأقليات الإسلامية

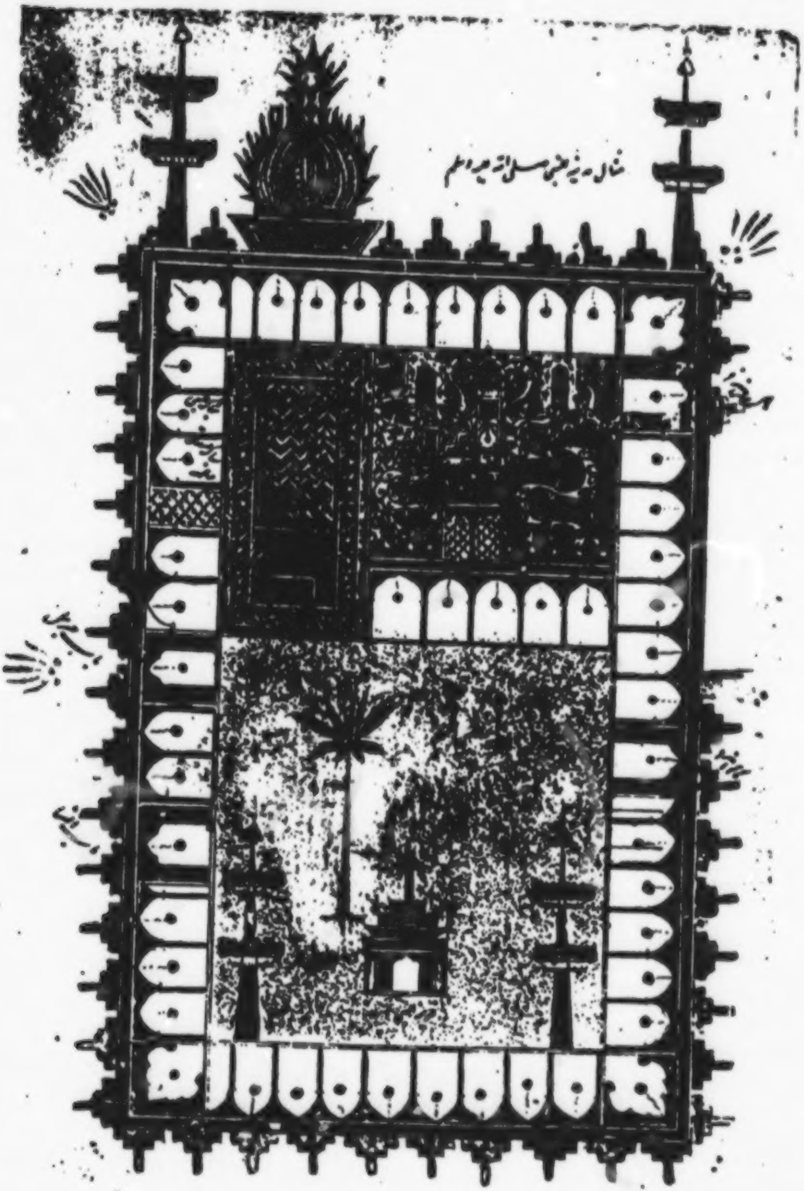
مركز الملك فيصل للبحوث والدراسات الإسلامية



Al-Khaṭṭ
Al-'Arabī
Min Khilāl
Al-Makhṭuṭāt

Arabic Calligraphy in Manuscripts

THE UNITY OF
ISLAMIC ART



AN EXHIBITION OF ISLAMIC ART AT THE ISLAMIC ART GALLERY
THE KING FAISAL CENTER FOR RESEARCH AND ISLAMIC STUDIES
RIYADH, SAUDI ARABIA, 1405 AH/1985 AD

Data Transferance Between CDS/ISIS and MINISIS

Presented by Xi Zhenyan
Written by Kin Zhenyan
Technical Services Office
Shanghai Population Information Centre

ABSTRACT:

The author discribes the procedure and experience in transferring data between CDS/ISIS and MINISIS. A great convenience is provided by MINISIS (G.00) for such transfers. But attention should still be paid to performance by users. Readers may get enlightenment in this paper, thus making their data tranfers go more smoothly.

Before MINISIS (G.00) was released, users had to run a programme written by themselves to rewrite the ISO file to make it acceptable for CDS/ISIS, or to revise the ISO.MST into a data format which could be processed by MINISIS. The data conversion between these two softwares was somewhat inconvenient then. An ISO file of CDS/ISIS type is now defined in the ISOC0NV processor of MINISIS (G.00). Only "CDS/ISIS" should be used to reply to the prompt "ISO TYPE" when users define a CD base. The ISOCONV will convert an ISO file which was exported from a CDS/ISIS data base into MINISIS or vice versa. No revision should be made in ISO file.

An HP Micro 3000 computer system was installed in December 1987 at Shanghai Population Information Centre (SPIC). Before then, there was only an IBM PC/XT in our centre. We created a bibliographic data base with CDS/ISIS (Mini-micro Version, Release 1.0) at the beginning of 1987 in order to check duplicatation in the holdings of our library. The structure of the data base was very simple. Each record consists of these eight fields:

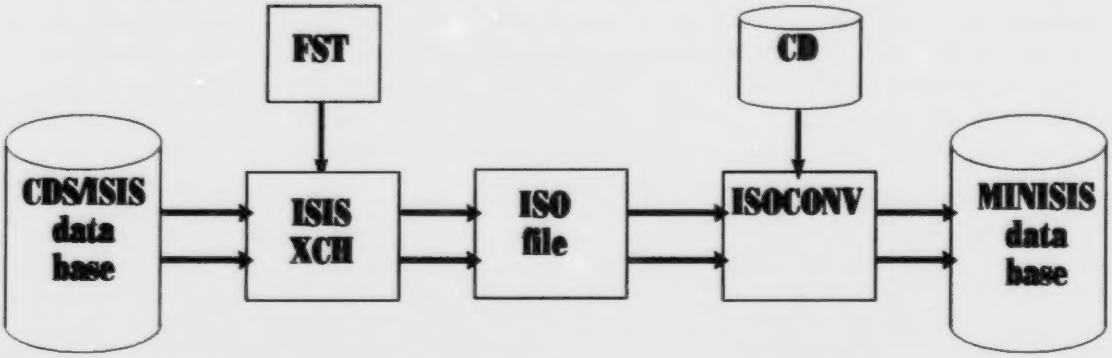
title,
author(s),
publication,
ISBN,
ISSN,
status,
call number and
note.

The "author(s)" field was repeatable and the "publication" field included three subfields: publisher, location of publisher and the date of publishing. The "status" field recorded if the book had been ordered but not received, or received, or received through exchanging. The "note" field recorded the way in which the book was ordered, the date when ordered, the date when received and the price of the book etc. The field definition table (FDT) of the data base is as follows:

```
Field Definition Table (FDT)                Data Base: SPICL
Tag  Name          Length  Type   Rep.  Delimiters/Pattern
10   TITLE            794    X
20   AUTHOR(S)         70    X      R
30   PUBLICATION      200    X            ABC
40   CALL NUMBER        7    X
50   STATUS             1    A
60   ISBN              13    X
70   ISSN               9    X
80   NOTE             201    X
```

MINISIS (G.00) was restored on our system in October 1987. After a MINISIS data base was created, all the 1294 records of CDS/ISIS data base were converted into MINISIS in order to put full use of the data in the CDS/ISIS data base, and to avoid unnecessary re-entry of the data, we practiced the performance of data conversion from MINISIS to CDS/ISIS and made preparations to communicate data to CDS/ISIS users.

The following illustration describes the overall course of our practice.



The data conversion of CDS/ISIS is realized with the ISISXCH programme. The field select table (FST) likes a switch, which can be used to select the fields that will be imported or exported and to change the tags of those fields. The data conversion of MINISIS is realised with the ISOCONV processor. CD data bases in MINISIS also like a switch, which can be used to change the tags of these fields. Data transfers between these two softwares are performed with the ISO file as an intermediary.

The procedure and experience of our practice is briefly described here.

Step 1. To export a data file of ISO 2709 format from CDS/ISIS.

1. The definition of a field select table (FST).

The ISISDEF programme of CDS/ISIS was invoked to define a reformatting FST named SPICLS. The "U" option was selected in the EXDEF menu. When the menu EXDBU was displayed, "E" was selected to create a FST. After entering the name of the FST (SPICLS), we began to define the fields which would be transferred. The following three points should be mentioned:

(1) The field's IDs in FST could be the same as, or they can differ from their original IDs in the data bases. The number of fields IDs in the FST could be equal to or less than the number in the data bases.

(2) Options 1,2,3 were not selected in our first FST which was used to invert the MASTER file of the data base. So indexing techniques (IT) 0 was used to index whole fields. This was done to all the fields selected for reformatting in the FST SPICLS. If IT 4 had been chosen in one field, the field separators # would have been seen between words in the field is ISO.MST, which had been exported from our data base. After ISO.MST was converted into a MINISIS data base, the field would contain only one word when it was repeatable, and only the first word would remain when it was in an elementary field. We did not test IT 1, 2, 3 in our reformatting FST SPICLS.

(3) The data extraction format of the elementary field used its tag to which V was prefixed. No other characters could be prefixed or the system could not find the field to be converted. For the repeatable fields, the data extraction format should be their tags to which V is prefixed and "/" is suffixed, and should be enclosed in parentheses, for example, (V200/). The reformatting FST we used is as follows:

```
Data Base Name: SPICL                    FST Name: SPICLS
 ID    IT    Data Extraction Format
 10    0     V10
 20    0     (V20/)
 30    0     V30
 40    0     V40
 50    0     V50
 60    0     V60
 70    0     V70
 80    0     V80
```

A data file of ISO 2709 format is produced with the reformatting FST shown above.

2. To export the MASTER file of a CDS/ISIS data base.

When running the ISISXCH program, "E" should be chosen in the EXCH menu. A worksheet will be displayed for data interchange. No revision was needed in the default values presented. We entered the data base name from which data would be exported, the range of records to export, and the reformatting FST name (SPICLS.) A file named ISO.MST with "#" as its field and record separator is produced after we exited from the EXCH menu. The ISO.MST file could be saved on either hard or soft disc, according to need.

Part of the records in ISO.MST are displayed here:

```
00210000000000097000450000000050006000000010005000006020001200056030003400068
040000102060000200110#00059#Population and Society in Twentieth Century
France#Dyerolin#^aHoddar and Stoughton Ltd^c1978#B100013#R##
0026500000000009700045000050006000000100053000060200014000590300084000730400
00157060000200165#00060#Research Management in Family Planning Programs, No.
^#Q, John S.#^aAsian and Pacific Development Administration Center, Kuala
Lampur, Malaysia^c1978#B100020#R##
0023500000000010900045000050006000000100063000060200012000690200011000810300
000920040000800115050000200123#00061#A Brief Introduction of Shanghai
Population Information Centre#Zheng Yiliu#He
Tiefeng#^aSPIC^bShanghai^c1987#B101010#R##
```

Let us take one record from this sample and analyze its structure:

```
0023500000000010900045000050006000000100063000060200012000690200011000810300
000920040000800115050000200123#00061#A Brief Introduction of Shanghai
Population Information Centre#Zheng Yiliu#He
Tiefeng#^aSPIC^bShanghai^c1987#B101010#R##
```

The first twenty-four bytes of the record are the LEADER, which is followed by the DIRECTORY and the DATA successively. This record consists of 7 fields. Field 3 is repeatable and occurs twice. Fields with no data do not appear.

In the DIRECTORY section, each field is described with 12 bytes. Take field 2 for example. Its DIRECTORY is as follows:

```
010006300006
```

The first 3 bytes are the tag of field 2. The following 4 bytes are the length of field 2 and the last 5 bytes are its starting character position. This is the value of the starting position of field 1 plus the length of field 1. The rest can be inferred.

There are two occurances of field 3 in the DIRECTORY because field 3 is repeatable.

Field 4 consists of 4 subfields. Its tag is "030." No special marks are found in the DIRECTORY for this field.

In the DATA section, we see the field separators # and record separator #. At the end of the record, ## is seen. The first # is a field separator and the second # is a record separator. The separators between repeatable fields are also #, as in:

Zhen Yiliu#He Tiefeng

Its repeatability has been indicated in the DIRECTORY entry. The field tagged 030 includes 3 subfields. Three CDS/ISIS subfield markers ^a, ^b, ^c can be seen in the DATA section.

Step 2. To transfer ISO.MST file to the HP Micro 3000.

The AdvanceLink software was used to transfer the ISO.MST file from disc on the PC to the HP Micro 3000. After pressing the function key for Transfer File, an AdvanceLink menu is displayed. In the menu, the transfer direction should be "to remote," the other default values in the menu may not need to be changed. Pressing the function key for continue brings up the next screen. This is where ISO.MST and its directory are entered next to the prompt for "Local File Name" and any file name which follows the MPE rules could be used in response to the "Remote File Name" prompt. The file transfer type should be standard ASCII and the record size is 80. Other default values remain and the transfer can be started by pressing the function key for continue.

Step 3. To create a MINISIS data base.

In order to successfully transfer data from a CDS/ISIS data base to a MINISIS data base both a MINISIS CD definition and an RD definition are required. There are several points that should be noted.

1) The length of fields in the RD data base should be longer than or equal to the corresponding field lengths in the CDS/ISIS data base. To create a MINISIS data base whose structure is similar to the CDS/ISIS data base, repeatable fields and subfields in

the MINISIS data base should correspond with those in the CDS/ISIS data base. Subfields could be changed to elementary fields through transfer if necessary.

2) When creating the CD definition, the default value in the ISO type of DB HEADER SCREEN is ISIS. Four ISO file types are supported for DUMP (ISO to MINISIS): ISIS, AGRIS, Micro CDS/ISIS, and UNISIST. ISIS, Micro CDS/ISIS and UNISIST are supported for LOAD (MINISIS to ISO). In our case, the default value of ISIS is changed to CDS/ISIS.

3) The number of fields created in the CD definition might be less than or equal to that in the reformatting FST of CDS/ISIS. The selection of fields to be transferred could be done again at this point.

4) In the CD definition, the identifiers for the CDS/ISIS fields should be identical to those in the reformatting FST of CDS/ISIS. The identifiers for the MINISIS fields should be identical to those in the MINISIS RD.

5) For fields with subfields, "Y" should be entered at the "SUBFIELD Y/N" prompt in the RD-MASTER FIELD TAG worksheet, and the subfield identifiers of CDS/ISIS (^a, ^b, ^c in our case) should be entered in the SUBFIELD IDENTIFIERS in the CD FIELD TAG worksheet.

6) For repeatable fields, "Y" should be entered at the REPEATABLE Y/N prompt in the RD-MASTER TAG worksheet.

Step 4. To transfer the ISO file into MINISIS.

1) The ISO file was transferred from the PC to the HP using AdvanceLink.

In our case, it was name SPICL. ISOCONV is written for tape input (DUMP) or output (LOAD) and searches the tape for a file named ISO. The file equation

```
:FILE ISO = SPICL;DEV = DISC
```

is required to identify our file for ISOCONV.

2.) The treatment of ISOLIST files.

Each time we run ISOCONV, it sends a listing of the correspondance table to the system line printer. The name of this output file is ISOLIST. ISOLIST may be redirected to another output device with a file equation like this

```
:FILE ISOLIST;DEV = <logical device number..ie. terminal>
```

3) Running the ISOCONV processor.

ISOCONV could be chosen from the MINISIS menu. We replied to the prompt to ENTER COMMAND TO SELECT DATA BASE with the following:

```
DUMP FROM = name of CD data base TO = name of RD data base
SUBSET = all
```

When we want to create an ISO file from a MINISIS data base, the DUMP command is replaced by the LOAD command.

```
LOAD FROM = name of RD data base TO = name of CD data base
```

The ISOCONV processor may be interrupted with a CTRL-Y to obtain a listing of the number of records processed.

4) The treatment of E226 error message.

The main problem we came across was the E226 error message during data conversion. The message is E226 DIRECTORY OVERFLOW. M609 WRITE ERROR ON INVERTED FILE may also appear at this time. The name of the inverted file should be noted.

We responded to the E226 message by

(1) choosing the INVERT processor from the MINISIS menu

(2) selecting TREEMANT to change the size fo the B-Tree inverted file.

The root name of the B-Tree file which was display in M609 should be entered at the prompt ENTER KEY FILENAME and we responded with N to the prompt RECOVER WASTE SPACE which allowed us to obtain the statistics on the B-Tree. We could find the expenditure of both DIRECTORY and POSTINGS, and could thus determine how to change the size of the B-Tree file. The size of the B-Tree file should not be enlarged too much in order to save on disc space. The enlargement only deals with the distribution of space and does not cause loss of data. So the enlargement of the size of a B-Tree file must be done step by step during the conversion.

(3) To continue the conversion.

Data could then be converted again. There is one thing to be noted: data which has already been converted should not be converted a second time. The range of the subset should reflect the number of records converted when the E226 error occurred.

:SUBSET = the range of records starting with the one immediately after the record converted before the E226 error occurred.

That completes the description of our methods for transferring data from CDS/ISIS to MINISIS. The data transfer in the other direction (MINISIS to CDS/ISIS) can be done in much the same way. Attention should be paid to the following two points:

1) If the output file of MINISIS is to be saved on disc, a file equation is required before ISOCONV is run.

```
:FILE ISO;DEV = DISC;DISC = n,32,1;REC = -80,1,V,ASCII
```

where n is the maximum number of logical records. The default value is 1023. This value must be greater than the number of records to be transferred.

2) There should be enough free space on the PC's hard disc. The ISO file is the output of a MASTER file in MINISIS. If must be inverted after being imported into CDS/ISIS for the the purpose of retrieval. The inversion of a CDS/ISIS MASTER file needs a large space. If the space available on disc is not large enough, the inversion will be suspended.

The ISOCONV processor of MINISIS can be run easily and is flexible. MINISIS (G.00) provides users iwth a very nice possibility of sharing resources with users of CDS/ISIS. We have practiced transferring data using English bibliographic data bases. We will now beginning experimenting with the transfer of Chinese bibliographic data bases. We hope to strengthen the contact, cooperation and communication with other users of MINISIS in data conversion in order to improve our ability in MINISIS.

References:

1. Manual of CDS/ISIS (Mini-micro Version, Release 1.0)
2. Manual of MINISIS (G.00)
3. General Information Manual MPE V Commands

CDS/ISIS与MINISIS之间的数据转换

上海人口情报中心 席臻衍

摘要：本文主要介绍著者在CDS/ISIS与MINISIS之间进行数据转换的具体步骤和每一步中遇到的问题以及解决的方法。G版MINISIS为这种转换提供了极大的方便。但是，用户在实践中还有一些值得注意的地方。读者或许能由本文得到一点启发，从而使转换工作能顺利完成。

G版以前的MINISIS用户要实现MINISIS与Micro CDS/ISIS之间的数据转换，必须运行一个自己编制的程序来改写由MINISIS输出的ISO文件，才能转入Micro CDS/ISIS；对CDS/ISIS输出的ISO.MST文件也必须改写成MINISIS能处理的数据格式，才能转入~~CD~~MINISIS。两个软件之间的数据转换

比较麻烦。在G版MINISIS的ISOCONV模块中，专门为CDS/ISIS定义了一种称为CDS/ISIS类型的ISO文件。用户只要在定义CD类型数据库时，对"ISO类型"提示回答"CDS/ISIS"，ISOCONV模块就能通过CD库转入一个从Micro-CDS/ISIS ~~输~~输出的ISO文件，或者输出一个能被Micro-CDS/ISIS接收的ISO文件。用户不必再对ISO文件作任何修改。

上海人口情报中心是1987年12月安装的HP MICRO 3000小型机。在此以前，仅有一台IBM PC/XT微型计算机。为了解决图书馆々藏英文图书的采购查重困难和进行一些简单的检索，于1987年初建立了一个Micro CDS/ISIS的书目数据库。该库结构比较简单。每条记录仅有8个字段即书名、著者、出版、ISBN、ISSN、状态、索书号和备注；其中著者字段为可重复字段，出版字段包含三个子字段即出版者、出版地、出版年；状态字段记载该书是"已订未到"还是"已到"还是"交换赠阅"三种状态；备注字段则记录该书的订购渠道、订购日期、订购时书价

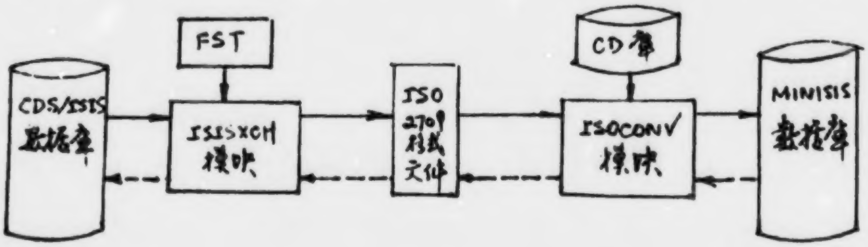
到书日期等内容。该库的字段定义表~~如~~（FDT）如下：

Field Definition Table (FDT)

Tag	Name	Len	Type	Rep	Delimiters/Pattern
10	TITLE	794	X		
20	AUTHOR(S)	70	X	R	
30	PUBLISH	200	X		abc
40	CALL NUMBER	7	X		
50	STATUS	1	A		
60	ISBN	13	X		
70	ISSN	9	X		
80	NOTE	201	X		

1988年10月我们安装了MINISIS（G版）。在建立MINISIS书目数据库时，为了利用原有CDS/ISIS数据库的数据，避免重复的输入劳动，我们于1989年初把CDS/ISIS库中的1294条记录全部转入MINISIS库中。同时，试运行了把数据从MINISIS数据库转换入CDS/ISIS库的过程，也为今后与CDS/ISIS用户进行数据交换准备了条件。

全部转换过程可用下~~边~~面的框图来表示：


FST
CD库
CDS/ISIS 数据库
ISISXCH 模块
ISO 2709 格式 文件
ISOCONV 模块
MINISIS 数据库

CDS/ISIS数据库通过ISIS XCH模块来实现数据的输入或输出，字段选择表（FST）就如一个转换开关，用户可以利用该表选择要转换的字段和改变字段标号。MINISIS则通过ISOCONV模块来实现数据的输入或输出，CD库也象一个转换开关，用户可以在CD库中选择要转换的字段和改变字段标号。两个数据库之间的数据转换是以格式为ISO 2709的文件为中间媒介。

下面将简单地叙述我们的具体做法和体会。

第1步，从CDS/ISIS数据库输出一个格式为ISO 2709的数据文件。

1．定义一个字段选择表（FST）

运行CDS/ISIS的ISISDEF模块，在EXDEF菜单中选择"u"修改数据库定义。在进入EXDBU菜单后选择"E"建立FST。这时，键入要建立的FST表名，例如splcls，就可以开始定义要转换的字段。在定义时要注意如下3点：

（1）表中的字段标识（ID）可以与原库的字段一样，也可以不一样。表中的字段数可以等于或少于原库的字段数。这就是说，可以根据

需要把原库主文件中部分或全部字段转换成ISO 2709格式的数据文件。

(2)由于我们原库中的用于倒排的字段选择表中未用到1、2、3种索引技术（IT），因此，我们用于输出的字段选择表里，每个字段的索引技术（IT）都选择"0"，即对全部字段作索引。如果选择"4"，则在输出文件中可以看到词与词之间全用"#"（字段分隔符）分隔，转入MINISIS后，在可重复字段里，每个字段仅含一个词；在基本字段中则丢失一部分数据，仅保留第一个词。关于第1、2、3种索引技术在输出文件中的效果我们未作试验。

(3)基本字段的抽词格式应为"V"加上字段标号如"V100"、"V300"等，不能用其他的字母来代替"V"，否则系统找不到要转换的字段。对于可重复字段来说，抽词格式为"V"加字段标号加斜杠，再用圆括号括起来。如(V200/)。我们定义的字段选择表名为spiclls，其参数列表如下：

Field Selection Table (FST)

ID	IT	Data extraction format
10	0	V10
20	0	(V20/)
30	0	V30
40	0	V40
50	0	V50
60	0	V60
70	0	V70
80	0	V80

用这个重新格式化参数文件输出的原库主文件就能符合ISO 2709格式。

2. 输出CDS/ISIS数据库的主文件。

运行CDS/ISIS的ISISXCH模块，在EXCH菜单中选择"E"(export)进入系统的输出工作单。对工作单里的预置参数一般不必作任何修改。只要填入原数据库名、要输出的记录范围和上述为输出主文件而定义的字段选择表名和重新格式化的参数文件名，如上述spic ks。在退出工作单时，系统就生成一个名为ISO.MST的数据文件。这个文件可以根据用户的要求存于A盘或B盘或C盘。

下面是选自ISO.MST文件的部分记录。

我们选取其中最后一条记录作为例子来分析一下该文件的结构（参见下图）。

```
00235000000000010900045000050006000000010006300006020001200069020001100081
<---------头标区 ------><-----------------------------------目次区 --------
                       <--字段  1--><--字段  2--><--字段  3--><--字段  3-->

030002300092040000800115060000200123#00061#A Brief Introduction of Shang
---------------------------------><------------------------------------
<--字段  4--><--字段  5--><--字段  6-->

hai Population Information Centre#Zhen Yiliu#He Tiefeng#^aSPIC^bShangh^c
------------------------------数据区 -----------------------------------

1987#B101010#r##
--------------->
```

记录的前24位字节是头标区，接着依次是目录区和数据区。该记录由7个字段组成，其中字段3重复1次，无数据的字段不出现。

在目录区内，每个字段用12位字节描述。以字段2为例，它的12位字节为：

010006300006

前3位是字段标号（TAG）"010"，第4~8位是字段长度"00063"，第9-12位是字段的起始位置"00006"。字段2的起始位置应等于字段1的起始位置加上字段1的长度，余类推。

由于标号为"020"的字段是可重复的，我们可以看到目录区里有2个字段3，它们的标号均为"020"。

字段4由3个子字段组成，标号为"030"。在目录区内无特殊标记。

在数据区内，我们可以看到字段分隔符和记录分隔符均为"#"。在一条记录的末尾既有最后一个字段的分隔符又有该记录的分隔符，因此，连续出现2个"#"。可重复字段之间的分隔符也是"#"，如上例中

Zhen Yiliu#He Tiefeng#

它~~们~~的该字段的可重复性已在目录区内记载。"030"是一个带子字段的字段，在数据区里该字段的位置上可以看到3个CDS/ISIS的子字段标识^a^b^c。

第2步，将ISO、MST文件送往HP3000。

运行advlink，请用进入相应的MINISIS的帐

户、组、用户，选择"传送文件"功能。

在"文件传送主菜单"中，"传送方向"应为"送至远地"。

"本地文件"应填入"ISO.MST"，并注明路径。

"远地文件"可由用户根据MPE操作系统的规则自行选择，設为SPICL。

"远地文件类型"应选"标准ASCII"。

"記录大小"为"80"。

其他可选项均用預置值即可。然后，"开始传送"。

第3步，建立MINISIS数据庫。

为了完成把CDS/ISIS数据庫的数据转换入MINISIS数据庫的工作，除了要定义一个RD庫外，还須定义一个CD庫，並且在建立MINISIS数据庫时要注意如下几点：

1)RD庫中的各字段长度应大于或等于相应的CDS/ISIS庫中的字段长；为了得到一个与CDS/ISIS庫结构相似的MINISIS庫，字段的可重复性与是否包含子字段均应与CDS/ISIS庫一一对应。

子字段在转换过程中可以根据需要转成基本字段。

2）定义CD库时，DB HEADER SCREEN中的ISO类型的预置值是ISIS。对于由ISO转入MINISIS的过程来说，可以选择ISIS、UNISIST、CDS/ISIS、AGRIS等四种不同类型的ISO文件；反过来，而对于由MINISIS转入ISO的过程来说，只有前三种可供选择。选择的类型不同，MINISIS的输出文件或对输入文件的要求自然也不同。在这里，我们应该把预置的ISIS类型改为CDS/ISIS。

3）CD库定义的字段数可以根据需要或小于或等于CDS/ISIS库输出的字段数。这里，我们经CDS/ISIS的字段选择表（FST）之后，对要转换的字段可以作第二次选择。

4）CD库中ISO字段标识符应与CDS/ISIS的FST（即上述例子中名为SPICLS的FST）中的标识符相对应。CD库中MINISIS的字段标识符必须与RD库相对应。

5）带子字段的字段在RD库的"RID-MASTER TAG"工作单中"SUBFIELD (Y/N)"处应填入"Y"，在"CD FIELD TAG"工作单的"SUBFIELD IDENTIFIERS"

处应填入CDS/ISIS数据库中所用的子字段标识如上例中的^a^b^c。

6)对于可重复字段，则在"RD-MASTER TAG"工作单的"REPEATABLE (Y/N)"处填入"y"即可。

第4步，转入MINISIS数据库

1)定义MINISIS与ISO文件的关系。

用advlink传送的文件存于HP 3000的系统的磁盘上。该文件在上述例子中定名为spicl。在运行MINISIS的ISOCONV模块之前，必须建立一个文件方程，把ISO文件引导到spicl文件。

```
:FILE ISO=SPICL;DEV=DISC
```

2)对ISOCONV输出文件的处理。

每运行一次ISOCONV模块，在系统的打印机上输出一份CD库的主要参数表。这个输出文件名为ISOLIST，可以用文件方程引导到其他的输出設备。例如，若要将其引导到当前操作的终端显示器上，则在运行ISOCONV之前建立下列文件方程：

```
:FILE ISOLIST;DEV=当前终端逻辑設备号
```

3)运行ISOCONV模块。

进入MINISIS，选择10(ISOCONV)模块。

```
>dump from=CD库名 to=RD库名
>subset=列出要转换的记录号范围或all(即全部记录)
```

系统即开始把ISO文件转入MINISIS。

若要从MINISIS中把主文件输出为ISO文件，则上述dump命令应改为：

```
>load from=RD库名 to=CD库名
```

在转换的过程中，同时按下ctrl和Y 2个键可以随时显示已经转换的记录数。

4)对E226出错信息的处理。

在向MINISIS库转入数据的过程中，我们遇到的主要问题是出現E226错误信息。

E226提示的信息是"目录文件溢出"。在出現E226的同时，也可能出現M609即倒排文件出错的信息。要注意M609提示的文件根名。

处理目录文件溢出的具体做法是：

(1)在MINISIS主菜单中，选择18(INVERSN)模块，

(2)运行TREEMANT，改变B-树倒排文件的大小。

输入M609出错信息中提示的B-树倒排文件的根名。先不要急于要求回收废弃的空间，在相应的提问处回答"N"。系统立即显示出该B-树文件的空间分配情况，目录部分和列表部分的用空间开销和尚余空间的大小。根据系统提供的这些信息，我们可以根据需要来改变其大小的方案。为节省磁盘空间，B树文件不宜一下子放得太大。好在这种放大仅涉及磁盘空间的分配不会丢失数据。

(3)继续转入数据.

放大B树文件后，可以再运行ISOCONV模块，继续数据转换过程。这时要注意的是不要把已经转入MINISIS的数据再转入，在MINISIS库中造成重复，为此

>SUBSET=自出现E226以后的记录号范围。

上面是我们把数据从CDS/ISIS库转入MINISIS库的全部过程。相反方向的转换在步骤上也相反。有两点需要指出：

1)在运行MINISIS的ISOCONV模块转换主文

件之~~前~~时，如果要将MINISIS的输出文件ISO存在磁盘上，则要先建立一个文件方程：

```
:FILE ISO;DEV=DISC;DISC=n,32,1;
 REC=-80,1,V,ASCII
```

其中n是的文件的逻辑记录条数，其预置值为1023，为使ISO文件能大到足以容纳MINISIS的输出文件，n应大于要转换的记录条数。

2) 在微机上要留出足够大的空间。

由advlink把ISO文件从HP 3000转入微机的仅是MINISIS数据库主文件的输出文件。由于CDS/ISIS没有相当于MINISIS联机倒排的功能，ISO文件输入CDS/ISIS库后还必须经过倒排才能进行检索。而CDS/ISIS在对主文件倒排时，其中间过程需要较大的空间。空间不够倒排就无法进行。

MINISIS的ISOCONV模块运行方便，选择灵活，它的G版本又为与CDS/ISIS用户共享资料创造了一个良好的条件。我们初步实现了英文书目数据库的数据转换，我们还将在汉化的

CDS/ISIS基础上实现中文书目数据库的转换。希望能与MINISIS的用户在数据转换方面加强联系，密切合作，互相交流，提高应用水平。

参考文献

1. CDS/ISIS (Mini-Micro Version, Release 1.0) 手册
2. MINISIS (G.00) 手册
3. General Information Manual MPE V Commands

An Experience in SPEEDWARE-MINISIS

Mr. Lic. Ricardo Gonzalez Moreno
Cooridinacion de Servicios de Computo
UNIVERSIDAD AUTONOMA METROPOLITANA–XOCHIMICO
MEXICO D.F.

In 1988, the Universidad Autonoma Metropolitana joined the MINISIS community with the idea of automating their three libraries.

The first part of this presentation will introduce the Universidad Autonoma Metropolitana.

The Universidad Autonoma Metropolitana (UAM) was created in 1974 by an official decree of the Mexican government and given the following objectives:

1) Education
2) Research
3) Cultural Diffusion

The UAM was basically created to be the University of Mexico City, and its physical installations are found around Mexico City. It has three university campuses called "units" and one general rectory. The units' names and locations are:

Unit name	Location
Azcapotzalco	North of city
Iztapalapa	East of city
Xochimilco	South of city

Each unit is independent with independent administrations and educational programs. Each unit has a capacity of 15,000 students.

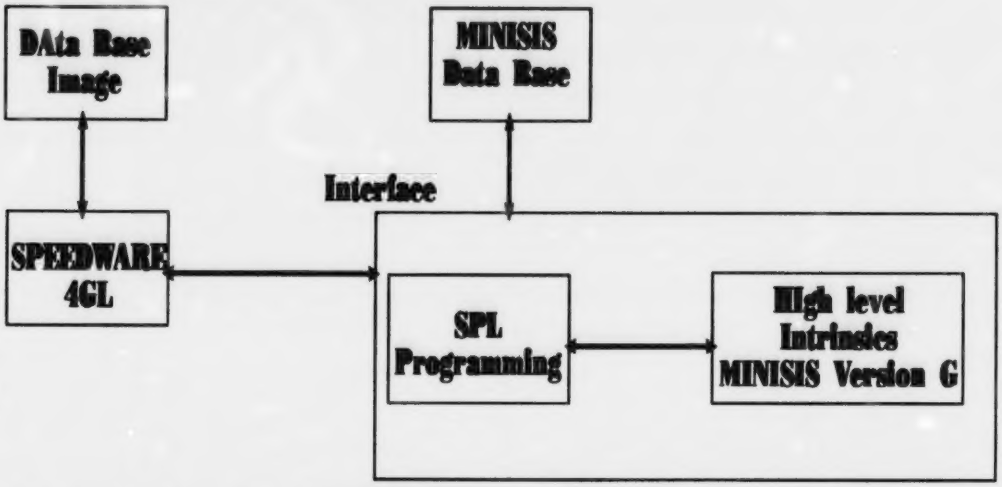
Each unit has its own Computer Center with a variety of computer makes and models. They are linked to a network called "TELEUAM." So we can connect from any unit to any computer in the network. We have 3 HP3000 series 52 computers, one in each unit.

In 1979 we began the development of the first library computing system and it was released in 1980, it is written in COBOL and FORTRAN and is installed in the Xochimilco unit. The limitations of this system, particularly the fixed field limit, prompted us to begin, in 1987, to begin the development of a new system. We were looking for a software for library information management and we found MINISIS to be the best selection.

We made the necessary moves to obtain the MINISIS software for our three installations.

At this point, some of you may be asking, "What is Speedware?" Speedware is a fourth generation programming language (4GL) used on HP3000 minicomputers and IBM PC compatibles. It is from Infocentre in Canada.

With Speedware you can develop Administrative systems in less time and they require less maintenance than those written in 3GLs such as COBOL, FORTRAN, PASCAL and C. With this language, we are creating an Administrative system and are planning to use it in our Library system which also involves administrative tasks. We studied the possibility of a Speedware program being able to access and obtain information from a MINISIS data base using an interface like this one:



This interface was developed in SPL. Tests were then made to prove that MINISIS data bases could be accessed from a SPEEDWARE application program.

At present, the UAM-Xochimilco has a computing system with which it has managed Circulation and Cataloguing for 9 years. Unfortunately, there is no relationship between them and there should be. The bibliographic system has to be able to satisfy our researchers' demands for information, so data has to be available from the second and third bibliographic levels. In addition to this, due to a growing demand, it is necessary to make a new integrated system which would include other areas such as acquisitions. To try to enhance the existing system is a complex and expensive task, as it is written in COBOL and FORTRAN and uses KSAM files.

The administration decided to create a new system.

The new system began with the MINISIS course given by Nicolas Cop from IDRC in the UAM-Xochimilco, Mexico.

After that, we realized that MINISIS was convenient for cataloguing, especially compared to other software products like SPEEDEX, IMAGE and KSAM among others. Also we realized that MINISIS could not manage the complex administrative procedures the new system required.

We attended the MUG held in Ottawa, Canada to get some experience about Administrative procedures using MINISIS and we found the following:

Washington's Library circulation system was done using SPL and the MINISIS low level intrinsics and this covered the practical aspects of the procedures.

IDRC's library covered acquistitions, cataloguing and searching with MINISIS and that they were going to implement circulation but their administrative procedures were practically all done manually.

In the CICH's system from the National University of Mexico "UNAM", the subscriptions system accesses MINISIS information through a file which is obtained by a constant batch update process. It is not run online.

We were acquainted during the MUG88 with a new generation of intrinsics for use in 3GL programs and we thought that they could be used in conjunction with the SPEEDWARE program.

An example of how to use the MINISIS high level intrinsics within a SPEEDWARE program is shown in Appendix 1.

The principal problems we found were:

> There is no detailed information on how SPEEDWARE works.
>
> Two technical problems with SPEEDWARE: it does not use variables greater than 256 bytes; and it does not have a variable type record.

It is important to take note that to be able to access MINISIS data bases from SPEEDWARE without the restrictions encountered and with the need for an interface, the SPEEDWARE software would require enhancements. It would be useful to ask for a new version of SPEEDWARE.

If we can program with SPEEDWARE, it would make our development simpler because the library's system could be online and directly accessed by the other systems such as accounting, budget and personnel.

Appendix 1

Abbreviated Forms:

SPEEDWARE PROGRAM:

```
Call OpenEnvironment;                               <routine SPL>
Accept "Tag Field: ",#tag;
Accept "ISN: ",#ISN;
Call ReadIsn (#isn,#stat);                          <routine SPL>
If #stat=ok then
     begin
          call GetField (#tag,#stat,#field);        <routine SPL>
          if #stat=ok then
                    print "Field value: ",#Field
     end
Call CloseEnvironment;                              <routine SPL>
```

SPL Routines

```
Procedure OpenEnvironment;
     Begin
          InitProg;
          Open Data Base;
          Open Dispositive;
     End.
Procedure ReadISN(isn,stat);
     Begin
          ReadDataBase(isn,stat);
     End.
```

```
Procedure GetField(tag,stat,field);
     Begin
          getfield(tag,stat,field);
     End.
Procedure CloseEnvironment;
     Begin
          Close Dispositive;
          Close Database;
          EndProg;
     End.
```

For additional information:

Mr. Ricardo González Moreno
Coordinador de Servicios de Cómputo
Mr. Víctor Hugo Hernández Torres
Jefe de Diseño de Sistemas
Universidad Autonoma Metropolitana - Xochimilco

Calz. Del Hueso 1100

Col. Villa Quietud, CP 04960, Mexico D.F.

Mexico

Tel: (5) 594-78-33 ext. 550 & 553

Fax: (5) 671-16-21

ADVANTAGES OF USING MINISIS IN THE DEVELOPMENT OF THE INFORMATION DATA BANKS FOR THE PROJECT: CULTURAL UNIVERSE OF THE STATE OF MEXICO

Josè Ricardo Montenegro G.
Unity of Informatics Services
I.M.C.

Introduction

One of the activities that has been considered to be of great importance in the present government administration of the State of Mexico is the culture. That is why they created the Instituto Mexiquense de Cultura in September of 1987.

The main objectives of this institute are:

- Cultural investigation and rescue
- Diffusion of the cultural resources
- Restoration and preservation of the colleges
- Classification and cataloguing of the bibliographic, museum holdings and document materials
- Program creation and execution of cultural activities
- Diffusion of the native languages
- Protection and utilization of the cultural sites
- Installation and administration of cultural houses in the 121 regions of the State of Mexico

Mexico is a country with a great historical and cultural richness and especially in the State of Mexico it is represented by various materials and human resources. These materials and people are distributed throughout Mexico. It is difficult to find a single region which does not have a site for cultural activities. There are also a great number of monuments, native ceremonial centers and archelogical zones. There are also 170 libraries in the State Net of Libraries. In the historical files there are a concentration of documents dating back as far as the time of the colonization of Mexico (the conquest of Mexico by Spain) that contain important sources of historical information.

It was considered necessary, given all this, to develop systems able to gather, organize and integrate all the information. This would provide a universal extension of the culture in the State of Mexico.

This led to the creation of the Instituto Mexiquense de Cultura's Unity of Informatics Services (U.S.I.) which has the main objective of analysis and elaboration of computerized systems oriented towards the generation of information banks that allow, on the one hand, support for the control and protection of the cultural resources, and on the other, gives access to the information for students, investigators and the public attending installations.

A detailed analysis of the cultural items of the State of Mexico led us to the design of an informatics model and the in the selection of software for the system, we decided that MINISIS was the most useful and able to meet the goals of development.

The Cultural Information System (C.I.S.)

The C.I.S. is a union of data bases associated with 3 general items that are affiliated with collecting cultural data:

- colleges
- people
- cultural sites

Colleges

Data bases with references to:

- Bibliography
- Museum collections
- Monuments
- Historical documents
- Native languages

In all the cases the data bases are accessed through data submodels (DS) which gives different views of the information. For example, in the bibliographic data base, there are different DSs for consulting, technical processes and acquisitions. It is similar in the case of museum collections; there is a DS for each type of collection (anthropology, history, modern art, popular culture, etc.)

Currently, these banks have more than 60,000 references.

The data bank of native languages was designed to handle the translations of five regional languages:

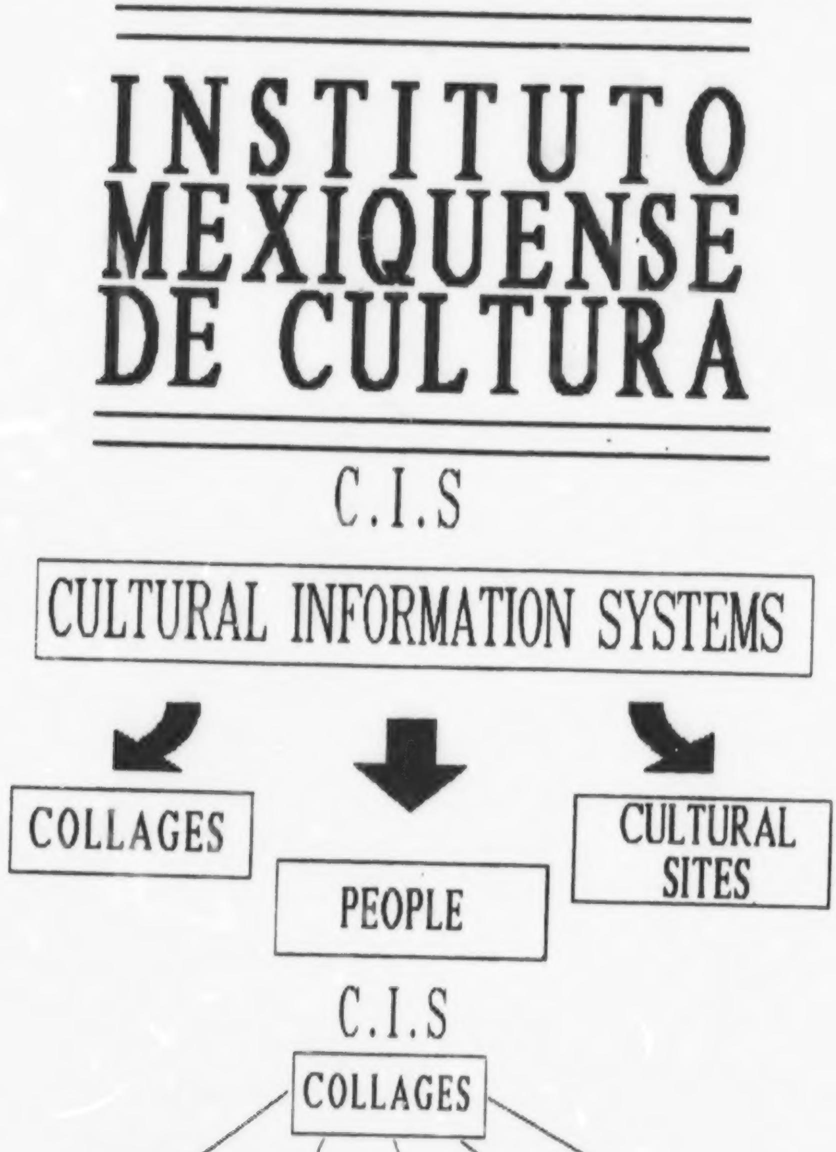
- Nahuatl
- Otomi
- Mazahua
- Tlahuica
- Matlazinca

There are more than 5,000 Nahuatl terms entered. These represent almost 16,000 references to Spanish translations. In its second stage, we plan to transfer the data base to CD ROM for distribution. The data bank of people for the culture is an integration of 3 groups:

- intellectuals
- artists
- illustrious men

The records are résumés or biographies (in the case of illustrious men) that make reference to the important people who have contributed in their activities to the nation.

One more of the C.I.S. data banks is the cultural sites. This bank allows us to know the details of the resources of the State of Mexico in relationship to the artistic and cultural activities programmed by the I.M.C. This data bank helps the organizers to select the appropriate site for events according to their character, the number of assistants required and the services required.


INSTITUTO
MEXIQUENSE
DE CULTURA
C.I.S
CULTURAL INFORMATION SYSTEMS
COLLAGES
PEOPLE
CULTURAL SITES
C.I.S
COLLAGES
BIBLIOGRAPHIC
MUSEOGRAPHIC
MONUMENTS
DOCUMENTAL
NATIVE MEXICAN LANGUAGES

C.I.S.

BIBLIOGRAPHIC

DATA-BASE DEFINITION

- DATE TYPE
- DATE-1
- DATE-2
- ILUSTRATIONS
- INTELECTUAL NIVEL
- REPRODUCTION FORM
- CONTENTS
- GUBERNAMENTAL PUBLICATION
- CONFERENCE?
- HOMANAGE?
- INDEX CONTAINS
- PRINCIPAL ENTRANCE
- FICTION?
- CLASIFICATION
- LANGUAGE
- CATALOGING
- DEWEY CODE
- ADQUISITION NUMBER
- I.S.B.N.
- PRINCIPAL SEAT
- PERSONAL
- CORPORATIVE
- CONFERENCE
- UNIFORM TITLE
- UNIFORM TITLE BY AUTHOR
- TITLE
- EDITION
- FOOTNOTE
- PHYSICAL DESCRIPTION
- TOPICS
- GENERAL NOTES

C.I.S.

MUSEOGRAPHIC AND MONUMENTS

DATA-BASE DEFINITION

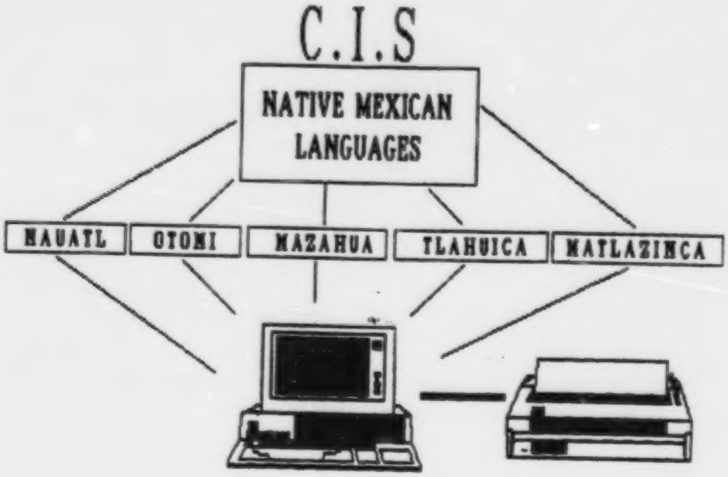
- INVENTARY NUMBER
- CATALOG NUMBER
- CLASIFICATION
- CODE OF CULTURAL SITE
- MUSEUM OR SITE LOCALIZATION
- GROUP
- OBJECT UBICATION
- MATERIALS
- REGION
- EPOQUE OR HISTORIC CICLE
- MANUFACTURE TECHNICS
- DECORATIVE TECHNICS
- CONSERVATION GRADE
- CULTURE (ORIGEN)
- AUTHOR
- TITLE OR DENOMINATION
- MEASURES
- NARRATIVE DESCRIPTION

C.I.S.

DOCUMENTAL

DATA-BASE DEFINITION

- CHRONOLOGYCAL
- DATE (FROM - TO)
- GEOGRAPHIC REFERENCE
- ORIGEN
- ORIGEN (COMPLEMENTARY)
- TEMATIC
- ABSTRACT
- INSTITUTION
- COLLAGE-CODE
- DESCRIPTIVE CODE
- DOCUMENT TYPE CODE
- INITIAL VOLUME
- LAST VOLUME
- EXPEDIENT
- PAGES
- INITIAL PAGE
- LAST PAGE

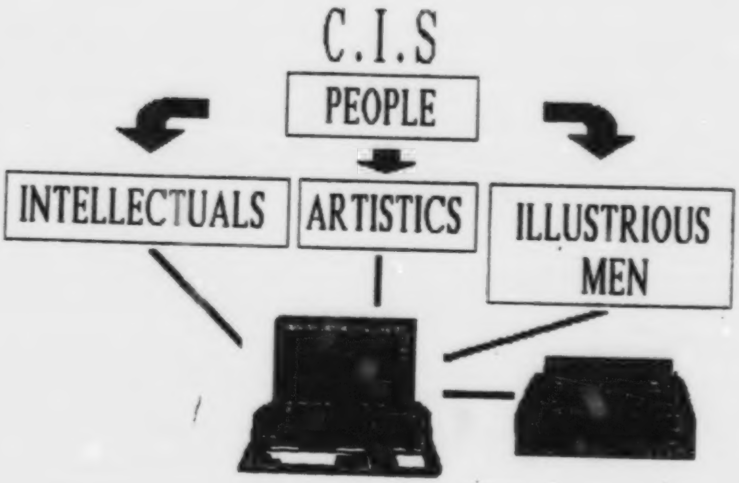

C.I.S
NATIVE MEXICAN LANGUAGES
NAUATL
OTOMI
MAZAHUA
TLAHUICA
MATLAZINCA

C.I.S.

NATIVE MEXICAN LANGUAGES

DATA-BASE DEFINITION

-NATIVE LANGUAGE CODE
-ALPHABETIC GROUP
-NUMERIC PREFIX WORD
-NATIVE WORD
-SPANISH WORD (TRANSLATION)
-NATIVE WORD DESCRIPTION
-NATIVE ETIMOLOGY


C.I.S
PEOPLE
INTELLECTUALS
ARTISTICS
ILLUSTRIOUS MEN

C.I.S.

PEOPLE

DATA-BASE DEFINITION

RECORD OF HUMAN RESOURCES

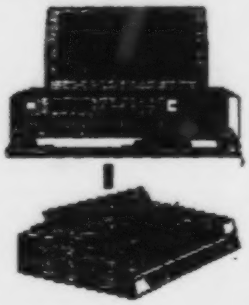
GROUP (I = INTELLECTUAL)
(A = ARTISTIC)
(H = ILLUSTRIOUS MEN)
NAME FIRST LAST
DENOMINATION
PSEUDONYMOUS
BIRTHDATE (DAY) (MONTH) (YEAR)
BIRTHPLACE
DECEASE-DATE
ADDRESS- STREET NUMBER (EXT) NUMBER (INT) TOWN/CITY
ZIP-CODE
TELEPHONE

PROFESSION
OCCUPATION
JOB TITLE
ACADEMIC DEGREE
ACTIVITIES OF HIS/HER JOB
NUMBER OF MEMBERS (IN GROUP CASE)
TITLES OF PRINCIPAL WORKS
EXHIBITIONS
MAIN AWARDS
PROFESSIONAL EXPERIENCE
OBSERVATIONS

C.I.S

CULTURAL SITES

MUSEUMS AUDITORIUMS
LIBRARIES THEATRES
ARCHEOLOGICAL ZONES
CULTURAL HOUSES
MONUMENTS
OTHER INSTITUTIONS



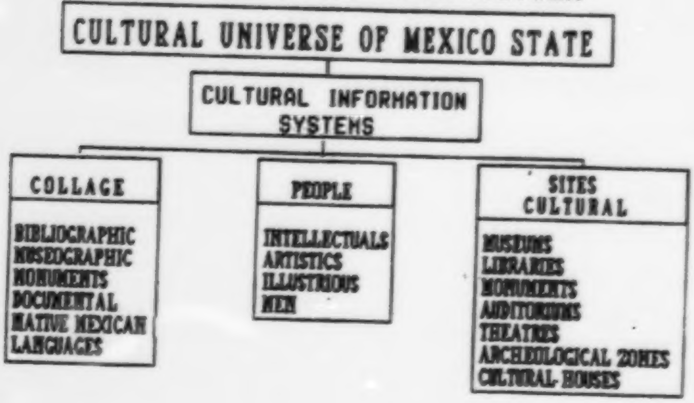
C.I.S.

CULTURAL SITES

DATA-BASE DEFINITION

CLASIFICATION : (M = MUSEUM) (C = CULTURE HOUSE) (B = LIBRARY)
(Z = ARCHEOLOGICAL ZONE) (T = THEATRE) (A = AUDITORIUM)
(N = NATIVE CEREMONIAL CTR)
LOCALIZATION
  DENOMINATION
  IN CHARGE OF
  GEOGRAPHIC UBICATION
  ADDRESS
  REGION
CAPACITIES
  TOTAL CAPACITY
  NUMBER OF SEATS
SERVICES
DIMENTIONS (A = GLOBAL) (B = OCCUPIED AREAS) (C = FREE AREAS)
AREAS CHARACTERISTIES
PROBABLE USES

UNITY OF INFORMATIC SERVICES OF INSTITUTO MEXIQUENSE DE CULTURA


CULTURAL UNIVERSE OF MEXICO STATE
CULTURAL INFORMATION SYSTEMS
COLLAGE
BIBLIOGRAPHIC
MUSEOGRAPHIC
MONUMENTS
DOCUMENTAL
NATIVE MEXICAN LANGUAGES
PEOPLE
INTELLECTUALS
ARTISTICS
ILLUSTRIOUS MEN
SITES CULTURAL
MUSEUMS
LIBRARIES
MONUMENTS
AUDITORIUMS
THEATRES
ARCHEOLOGICAL ZONES
CULTURAL HOUSES

Multilingual Thesaurus Using MINISIS

Sun Minglei
Zhou Jiyu
Shanghai Population Information Center

Abstract: There are two parts in this paper. One discusses the creation and application of the thesaurus which is used in the MINISIS data base, and the problems which should be noted. Some examples are given. The other part deals with some special problems we encountered using bilingual (Chinese/English) terms in the MINISIS thesaurus.

Classification and thesaurus are two methods used in information retrieval. The system structure and methodology of classification restricts multiple and detailed approaches to the documents. With the development of science and technology, and the changing contents of documents, the thesaurus serves better and has been adopted more often than the classification method.

MINISIS is an information management system which provides users with the capability to manage as many as 9 different languages in a database. In this paper we would like to discuss our use of the MINISIS thesaurus in two parts. The first describes the creation of our MINISIS thesaurus, and the second discusses some special problems encountered using a bilingual (Chinese/English) thesaurus.

Part I. Creation and Application of the MINISIS Thesaurus.

The process of creating and applying a thesaurus is briefly illustrated as follows:

a. Create a database (our example is named DBA) for entering the thesaurus terms.

b. Enter the validation terms into the database.

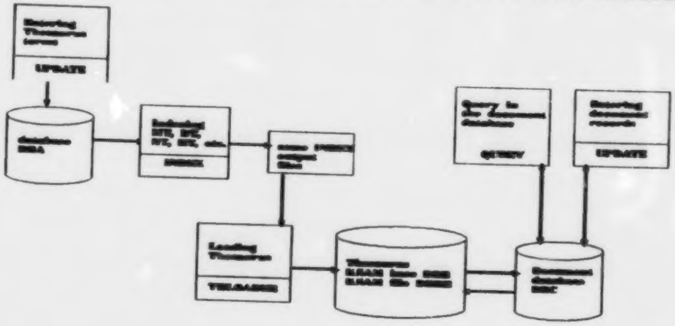
c. Create a second database (our example is named DBB) for loading the thesaurus. It will be a KSAM THES file. DBB will be used as the thesaurus inverted file.

d. Using THLOADER, the terms of DBA are loaded into DBB with index output files. The thesaurus is created.

f. Create a document database (our example is named DBC.) It is a MASTER type database. One or more fields in the database can use the thesaurus.

g. The data in the fields using the thesaurus are validated by the thesaurus as they are entered and accessed through the thesaurus for retrieval.

The process and the relationships between the databases is illustrated here:



(1) Using DATADICT, DBA which will be loaded with the thesaurus terms and their relevant explanations is defined. In DBA's header section, the database is defined as an RD of type MASTER. The fields required and their features are listed here:

FIELD NAME	MNEMONIC	TAG	OFFSET	LENGTH	REPEAT	PROMPT
MAIN TERM-ENGL	MTE	A010	-1	40		Y
MAIN TERM-CHIN	MTC	A020	-1	40		Y
BROADER TERM	BT	A030	-1	40	Y	Y
NARROWER TERM	NT	A040	-1	40	Y	Y
RELATED TERM	RT	A040	-1	40	Y	Y
ANY TERM	ANY	A060	-1	40	Y	Y
FORBIDDEN TERM-ENGL	FTE	A070	-1	40	Y	Y
FORBIDDEN TERM-CHIN	FTC	A080	-1	40		Y
USE TERM-ENGL	UTEE	A090	-1	40		Y
USE TERM-CHIN	USEC	A100	-1	40		Y
FACET NUMBER	FACET	A110	-1	10		
SCOPE NOTE-ENGL	SNE	A120	-1	600		
SCOPE NOTE-CHIN	SNC	A130	-1	600		
HISTORY NOTE-ENGL	HNE	A140	-1	300		
HISTORY NOTE-CHIN	HNC	A150	-1	300		

MAIN terms are essential to any thesaurus, while other terms are selected according to need. MAIN terms, STRUCTURE terms, ANY terms, FORBIDDEN terms and USE terms (A010 – A100) are loaded into the KSAM database, DBB, to create a thesaurus which can be used by the document database DBC. FACET number and SCOPE NOTE are not loaded into the KSAM database and are used only to provide explanations.

(2) Using UPDATE, the thesaurus is entered into the database DBA. The loading of the thesaurus will fail if the data does not fit the structure of a thesaurus, or if the data does not conform to the MINISIS rules. The STRUCTURE terms (BT, NT, RT) exist for a term and the term must exist in the thesaurus as a MAIN term. ANY terms should also exist as MAIN terms. The following two points are also very important: 1. Each

FORBIDDEN term with its USE term should be in a separate record. 2. The USE term should exist as a MAIN term in another record.

Here is an example of a entry structure using the database definition from above which can be accepted by MINISIS:

```
ISN=1
A010 MTE        :cereals
A020 MTC        :谷物
A030 BT         :food
A040 NT         :wheat
                 rice
A050 RT         :agriculture
                 harvesting
A060 ANY        :agricultural production

ISN=2
A010 MTE        :food
A020 MTC        :食品
a040 NT         :cereals

ISN=3
A010 MTE        :wheat
A020 MTC        :小麦
A030 BT         :cereals

ISN=4
A010 MTE        :rice
A020 MTC        :稻米
A030 BT         :cereals

ISN=5
A010 MTE        :agriculture
A020 MTC        :农业
A050 RT         :cereals
                 harvesting

ISN=6
A010 MTE        :harvesting
A020 MTC        :收获
A050 RT         :cereals
                 agriculture

ISN=7
A010 MTE        :agricultural production
A020 MTC        :农产品

ISN=8
A070 FTE        :cereal
A080 USEE       :cereals
A090 FTC        :谷子
A100 USEC       :谷物
```

As per the rule about records for each STRUCTURE term existing as a MAIN term in a record of its own, records 2 through 7 are necessary because of the contents of record 1. "FOOD" is a BT of "CEREALS" in record 1. So there must be a record 2, in which "FOOD" is a main term with the NT "CEREALS". "WHEAT" is a NT of "CEREALS". So there must be a record 3 with "WHEAT" as a MAIN term and "CEREALS" as a BT. And so on, with the rest of the records.

The relationships between ANY terms are quite loose. They are not organized to follow the logical division. MAIN terms which are not used often in searches can be grouped in an ANY table. The MAIN term should point the ANY term and the other STRUCTURE terms should point to each other. ANY term fields are repeatable as one MAIN term may have many ANY terms. One MAIN term may belong to many ANY tables. Different MAIN terms can point to the same ANY grouping in an ANY table.

(3) In INDEX, DBA is indexed to get index output files for the THLOADER process. Each MAIN, BT, NT, RT and ANY term has an index output file. The fields not used in the thesaurus do not need to be indexed. In our example we get 7 index files for DBA. If there are 1000 records in DBA and we decide to name the index output files sequentially as THMT, (thesaurus main term), THBT, etc. our session would look something like this:

For the MAIN terms:

```
ISN=1/1000
OUTPUT=THMT
KEY=MT,LENGTH=41,UPSHIFT=YES
FIELD=A010,SUFFIX="E"
FIELD=A020,SUFFIC="C"
END
***
```

For the BROADER terms:

```
ISN=1/1000
OUTPUT=THBT
KEY=MTERM,LENGTH=40,UPSHIFT=YES
FIELD=A010
END
KEY=BT,LENGTH=40,UPSHIFT=YES
FIELD=A030,BLANK=NO,OCC=100
END
***
```

For the NARROWER term and the RELATED term, we would follow suit with the entries for BROADER terms.

For the ANY terms:

```
ISN=1/1000
OUTPUT=THANY
KEY=ANY,LENGTH=40,UPSHIFT=YES
FIELD=A060
END
KEY=MTERM,LENGTH=40,UPSHIFT=YES
FIELD=A010
END
***
```

And for the FORBIDDEN-USE terms of the first language:

```
ISN=1/1000
OUTPUT=THFT
KEY=FTE,LENGTH=40,UPSHIFT=YES
```

```
FIELD = A070
END
KEY = USEE,LENGTH = 40,UPSHIFT = YES
END
***
```

The second language output can be created in the same way except for "OUTPUT = THFT1."

The keylength of the MAIN term index must be equal to the length of the field in the KSAM database DBB plus 1. This allows the suffix of the CSIC to sort easily. The language code in our case is "E" or "C." This code will be discussed more in step 5. There is a rule for the FORBIDDEN-USE output file name: the output file name of the first language FORBIDDEN-USE file should be THFT. The output file of the second language FORBIDDEN-USE file should be THFT1. Do not assign other names to files. UPSHIFT must be YES, otherwise all the terms in DBA must be entered in all capital letters. BLANK = NO suppresses those keys where the primary key but no secondary key exist. OCC = 100 allows the repeatable fields to occur up to 100 times in a record.

These INDEX commands are used repeatedly during the testing procedures. To avoid having to re-enter the commands manually, we replied YES to the prompt "DO YOU WISH TO SAVE KEY DEFINITION?" and giving a filename to the prompt for "PLEASE SPECIFY FILE NAME?" The file can be used in subsequent tests and modifications of the thesaurus. If, for example, the INDEX definition for the MAIN term indexing is saved in a file called MTFILE, the session for INDEXing the MAIN term would look like this:

```
ISN = 1/1000
OUTPUT = THMT
USE = MTFILE
```

(4) Using DATADICT, we defined DBB as a KSAM database into which we loaded the thesaurus which would be used by the document database.

To define DBB's header information we selected database type: RD, and file type: THES. The KSAM file name may be any four characters. Suppose DBBR is selected as the root file name for the inverted file, then DBBR must also be selected as the thesaurus root file name as well. After answering the key length and number of languages used, all fields are defined automatically. The definitions can be seen using the LIST command or the SHOW command in the subcommand handler menu. For our example, DBB with a term length of 40 characters and 2 languages a KSAM file called DBBR will be built. Each record will have the following fields defined:

FIELD NAME
TAG
LENGTH
MAIN TERM
X010
40
CN OF MAIN TERM

```
X020
6
CN OF LANGUAGE1
X030
6
CN OF LANGUAGE2
X040
6
ISN OF TERM
X050
6
```

X010 is the primary key of this file and is used in loading the terms of the thesaurus. X020 is the secondary key. CN is concept number. MINISIS translates, searches and establishes FORBIDDEN-USE terms using the CN.

(5) The thesaurus can now be loaded. For our example, the THLOADER session would look this:

```
ENTER NAME OF THESAURUS KSAM DATA BASE OR EXIT – DBB
FILE THESAURUS = THMT
FILE BT = THBT
FILE NT = THNT
FILE RT = THRT
FILE ANY = THANY
FILE FT = THFT
LANGUAGE EN = "E",CH = "C"
END
```

The first prompt is for the KSAM database name. The answer to the FILE prompts are the index output file names. "FILE FT = " should be the index output file for the first language. "E" and "C" in the language codes should be the same as the suffixes used to index the MAIN terms.

After these steps, the following things will be done by MINISIS automatically: assign a concept number to every term for translating; point from each FORBIDDEN term to the appropriate USE term; and build a structure file for loading the structure code and the concept number. Now the multilingual thesaurus DBB is created. During the creation of the thesaurus if illegal terms are encountered, the ISN and an error message are displayed. Use this information to correct the problem; this requires you to return to step 2, modify the information and then repeat steps 3 and 5.

The database DBB cannot be entered and modified directly. The contents are loaded through THLOADER. If the thesaurus requires modifications, modify DBA and repeat step 3 to index the relevant fields and step 5 to load the files into DBB.

(6) In DATADICT we can define DBC. The function of its fields will be controlled by the thesaurus should be:

VALIDATION (Y/N) – Y

ROOT NAME OF VALIDATION FILE – KSAM file name (DBBR in our example)

OFFSET TO FIELD – 4

LENGTH OF KEY TO EXTRACT – length of KSAM primary key field (40 in our example)

TYPE OF EXTRACTION (T/W/) – term or whole field

TYPE OF INVERTED FILE (K/B/) – K

ROOT NAME OF INVERTED FILE – KSAM file name (DBBR in our example)

THESAURUS INVERTED FILE (Y/N) – Y

(7) The validation field of the records in the document database will be controlled by the thesaurus. The characters in the validation field should be capitalized; only the MAIN term in the thesaurus will be accepted. Illegal terms, including FORBIDDEN-USE terms, can not be entered into this field. If an illegal term is entered, the system will show some legal terms, alphabetically, for selection.

We can query the thesaurus controlled field in the document database using different methods. MINISIS has a strong thesaurus function. Users can query for a term in one language and MINISIS can select records with the term in the second language. Queries can use the BT, NT, or RT notation in a search. A query for ANY terms selects records with one or more terms of the corresponding ANY table values. If the query is for a FORBIDDEN term, MINISIS uses its USE term automatically.

For example, using QUERY on the validation field could look like this:

```
>=cereals
   CEREALS    P=2
   谷物    P=1
     1:      P=3          T=3

Q> bt cereals
   CEREALS    P=2
   谷物    P=1
   FOOD    P=1
     2:      P=3          T=3

Q> nt cereals
   CEREALS    P=2
   谷物    P=1
   RICE    P=1
   稻米    P=2
   WHEAT    P=1
   小麦    P=1
     3:      P=7          T=7

Q> rt cereals
   CEREALS    P=2
   谷物    P=1
   AGRICULTURE      P=2
   HARVESTING     P=1
     4:      P=5          T=5

Q> any agricultural production
   AGRICULTURAL PRODUCTION    P=1
   农产品    P=2
   CEREALS    P=2
  谷物    P=1
    5:      P=6          T=6

Q> ft on
Q> cereal
   CEREALS
   谷物    P=1
     6:      P=3          T=3
```

If the query term is Chinese, the hit/records is the same. The difference in the results is simply that MINISIS gives the Chinese followed by the English hits.

In QUERY the terms of the thesaurus can be seen using the "DISFLAY [tag] @" command. FORBIDDEN terms can be seen using the "DISPLAY [tag] |@ command."

Part B. Some special problems encountered in compiling a bilingual English-Chinese thesaurus.

MINISIS is strong on multilingual thesaurus functions, including Chinese. So it is easy to create a bilingual English-Chinese thesaurus. Before entering every term in MINISIS, a written thesaurus should be made. There is no ready-made population thesaurus available now. Based on the English edition "Population Multilingual Thesaurus" and the POPLINE thesaurus, we compiled our own thesaurus. Because of a large gap between the Chinese and English translations, one English term may end up with more than one corresponding Chinese term. More than one English term may have one Chinese translation. We dealt with this in this manner:

(1) If the number of English FORBIDDEN terms is not the same as that of Chinese FORBIDDEN terms, they can be entered separately because the FORBIDDEN term field is repeatable. This results in one English term having one Chinese term. For example, the English MAIN term "bicycle" has a FORBIDDEN term "bike" and the Chinese MAIN term has two FORBIDDEN terms. So we handled it like this:

```
ISN=9
A010 MTE          : bicycle
A020 MTC          : 自 行 车

ISN=10
A070 FTE          : bike
A080 USEE         : bicycle
A090 FTC          : 脚 踏 车
                    单 车
A100 USEC         : 自 行 车
```

(2) There is one special case for (1). When one English MAIN(USE) term has one or more FORBIDDEN terms, but they have only one corresponding Chinese term. MINISIS can deal with it like this (using our case as an example): The corresponding Chinese term is entered in MTC, no more Chinese terms are entered in FTC or USEC. The translation function is not weakened. For example:

```
ISN=11
A010 MTE          : aging
A020 MTC          : 老 龄 化

ISN=12
A070 FTE          : ageing
A080 USEE         : aging
```

(3) The opposite problem is one where there is one English MAIN term which has no FORBIDDEN terms itself, but the corresponding Chinese term has one or more FORBIDDEN terms. This, like (2) can be dealt with easily. For example:

```
ISN=13
A010 MTE          : seal
A020 MTC          : 图 章

ISN=14
A090 FTC          : 印 鉴
                    印 章
A100 USEC         : 图 章
```

All these problems will appear in the translation between Chinese and English. The important key to solving these problems lies in the repeatability of the FORBIDDEN term field.

We practiced the creation of MINISIS thesaurus structures. Please be kind enough to give us your opinion.

MINISIS 中多种语言词表功能的使用

Multilingual Thesaurus Using in MINISIS

上海人口情报中心　　孙明磊　周基玉
Shanghai Population Information Centre
Sun Minglei, Zhou Jiyu

【内容提要】本文第一部分讨论在 MINISIS 中建立、调用词表的总体过程和详细步骤及每一步要注意的问题，并举例说明．第二部分讨论一词多译和多词一译的问题在 MINISIS 中英文词表中的处理方法．

【Abstract】 There are two parts this paper. One discusses the creation and application of the thesaurus which is used in MINISIS data base, and the problems which should be noted with some examples. the other deals with some special problems on using bilingual terms (Chinese and English) in MINISIS thesaurus.

分类法和主题法是图书情报资料工作中的两种不同的检索方法．分类法的体系结构和方法技术使其在对图书情报资料进行多元性和专指性检索时受到限制，因此，随着科学技术的发展和图书资料的内容的变化，分类法已难以适应用户对情报资料检索的不同要求．主题法则克服了分类法的短处，顺应了现代科学技术发展的趋势．

主题词表是主题法的基础．MINISIS 作为一个文献数据库管理系统，能进行九种不同语言词表的互译，使多语种文献的管理能在同一数据库内进行．本文根据我们的实践，分两部分讨论 MINISIS 中词表功能的运用．第一部分是 MINISIS 中词表的建立和调用；第二部分讨论如何在 MINISIS 中满足中、英文词表的一些特殊要求．

一、MINISIS 中词表的建立和调用

MINISIS 中词表的建立和调用，概括地说，可按下列步骤进行：

(1) 建立一个用于输入词表内容的数据库，设为 DBA，文件类型是 MASTER．

(2) 将符合要求的词表内容输入到 DBA 中．

(3) 对 DBA 做索引，得到若干个索引输出文件．

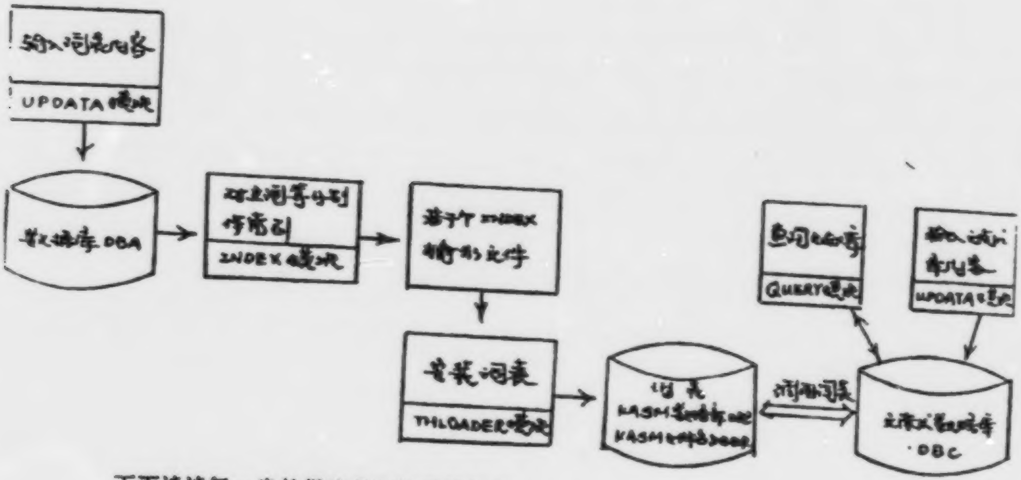
(4) 建立一个用于装词表的数据库，文件类型是 THES，属于 KASM 文件．它将作为文献数据库的词表倒排文件，也就是最后建成供文献库调用的词表，设为 DBB．

(5) 在词表装人模块下，通过索引输出文件将 DBA 中的内容装人 DBB，词表建成．

(6) 建立文献数据库，设为 DBC，文件类型是 MASTER．定义该库时，可选择其中一个或多个字段调用词表．

(7) 文献库 DBC 的录入，检索，其中已定义的词表倒排字段将受到词表的控制．

上述具体过程及各数据库之间的关系可用下图表示：


输入词表内容
UPDATA模块
数据库 DBA
对主词等分别作索引
INDEX模块
若干个INDEX输出文件
安装词表
THLOADER模块
词表
调用词表
文献数据库
·DBC
查词和检索
QUERY模块
输入文献内容
UPDATA模块

下面谈谈每一步的做法及要注意的问题.

(1) 在 DATADICT 模块下，定义数据库 DBA，这个库用于录人全部词表内容及相应的说明。该库头标定义中，库类型为 RD，文件类型为 MASTER，所需的字段及其性质，以中英文两种语言举例如下:

字段名	助记名	TAG	OFFSET	长度	重复性	提示
英文主词	MTE	A010	-1	40		Y
中文主词	MTC	A020	-1	40		Y
上位词	BT	A030	-1	40	Y	Y
下位词	NT	A040	-1	40	Y	Y
关系词	RT	A050	-1	40	Y	Y
ANY词	ANY	A060	-1	40	Y	Y
禁用词（英）	FTE	A070	-1	40	Y	Y
使用词（英）	USEE	A080	-1	40		Y
禁用词（中）	FTC	A090	-1	40	Y	Y
使用词（中）	USEC	A100	-1	40		Y
编号	FACET	A110	-1	10		
范围说明（英）	SNE	A120	-1	600		
范围说明（中）	SNC	A130	-1	600		
历史说明（英）	HNE	A140	-1	300		
历史说明（中）	HNC	A150	-1	300		

其中除主词字段是任何词表数据库所必备的之外，实际还需哪些字段可根据词表的不同要求而定，亦即其他字段都是可选的。主词、结构词、ANY 词、禁用词和使用词等字段（A010–A100）的内容将装人 KSAM 库 DBB，形成可被文献数据库 DBC 调用的词表。编号、说明字段（A110–A150）的内容不装人 KSAM 库，仅用于在 DBA 库中对相

应主词提供说明．

(2) 在 UPDATE 模块下，将词表的具体内容输入 DBA 中．录入的内容，不仅要符合一个词表应具有的结构关系，而且也要遵循 MINISIS 规则，否则在装入词表时将中途出错而导致失败．能被 MINISIS 接受的词表应该是：每个结构词（BT、NT、RT）、ANY 词必须在其他记录中作为主词存在、同时具有相应的结构词；禁用词和使用词的互相指代必须单独成记录，不能与主词（使用词）及结构词存在于同一记录中，而且使用词必须在其他记录中作为主词存在．

下面是一个可被 MINISIS 接受的词表例子（沿用步骤（1）中定义的字段）：

```
ISN=1
A010 MTE     :cereals
A020 MTC     :谷 物
A030 BT      :food
A040 NT      :wheat
              rice
A050 RT      :agriculture
              harvesting
A060 ANY     :agricultural production

ISN=2
A010 MTE     :food
A020 MTC     :食 品
a040 NT      :cereals

ISN=3
A010 MTE     :wheat
A020 MTC     :小 麦
A030 BT      :cereals

ISN=4
A010 MTE     :rice
A020 MTC     :稻 米
A030 BT      :cereals

ISN=5
A010 MTE     :agriculture
A020 MTC     :农 业
A050 RT      :cereals
              harvesting

ISN=6
A010 MTE     :harvesting
A020 MTC     :收 获
A050 RT      :cereals
              agriculture

ISN=7
A010 MTE     :agricultural production
A020 MTC     :农 产 品

ISN=8
A070 FTE     :cereal
A080 USEE    :cereals
A090 FTC     :谷 子
A100 USEC    :谷 物
```

如前面所述的规则：记录 1 的组成，使记录 2 至记录 7 必不可少。即由于记录 1 用“food”作为“cereals”的上位词，则必须有以“food”作为主词存在的记录 2，其中有相应的下位词“cereals”。又由于记录 1 用“wheat”作为“cereals”的下位词，则必须有以“wheat”作为主词存在的记录 3，其中有相应的上位词“cereals”。其他记录以此类推。反映用代关系的记录 8 必须单独成记录，不可并入记录 1 中。记录 8 中的使用词“cereals；谷物”必须有作为主词形式存在的记录 1。各个记录象同一样按上述规则，编制成了供使用的词表。

ANY 词，从词表结构意义上说，比较宽松，不必考虑严格的概念区分，可根据需要把存在某种关联的词汇集起来以方便检索。ANY 词在词表中单向指引，而不象其他结构词那样必须互相参照。从步骤（1）字段定义中可以看到，ANY 词被定义成可重复的，即一个主词可有多个 ANY 词，也就是可属于多个 ANY 表。同时，多个记录的主词可具有同一 ANY 词，构成该 ANY 词的 ANY 表。

（3）在 INDEX 模块下，对 DBA 做索引，以获得索引输出文件供装词表(THLOABER）用。对词表中的主词、上位词、下位词、关系词、ANY 词，每种语言的用代关系要各形成一个索引输出文件。词表中未用到的结构当然不用做索引。沿用上例，就需要形成七个索引输出文件。假设 DBB 中有 1000 条记录，索引输出文件名分别用 THMT，THBT 等表示，具体做法如下：

主词：

```
ISN=1/1000
OUTPUT=THMT
KEY=MT,LENGTH=41,UPSHIFT=YES
FIELD=A010,SUFFIX="E"
FIELD=A020,SUFFIX="C"
END
• • •
```

上位词：

```
ISN=1/1000
OUTPUT=THBT
KEY=MTERM,LENGTH=40,UPSHIFT=YES
FIELD=A010
END
KEY=BT,LENGTH=40,UPSHIFT=YES
FIELD=A030,BLANK=NO,OCC=100
END
• • •
```

下位词、关系词的做法与上位词类似。

ANY 词：

```
ISN=1/1000
OUTPUT=THANY
KEY=ANY,LENGTH=40,UPSHIFT=YES
FIELD=A060
END
KEY=MTERM,LENGTH=40,UPSHIFT=YES
FIELD=A010
END
• • •
```

第一种语言用代关系：

```
ISN=1/1000
OUTPUT=THFT
```

```
KEY=FTE,LENGTH=40,UPSHIFT=YES
FIELD=A070
END
KEY=USEE,LENGTH=40,UPSHIFT=YES
FIELD=A080
END
* * *
```

第二种语言用代关系的做法与第一种语言类似，其中"OUTPUT=THFT1"。

主词输出文件中的长度应是 KSAM 库 DBB 主键的长度加一，这一位用于给每个词加标识语种的后缀，以便排序。上例中为"E"或"C"。此后缀在下面步骤（5）中还要提到。用代关系的输出文件名有一个规定，例中第一种语言用代关系的输出文件名为 THFT，第二种语言的为 THFT1，而用 THFT、THFT2 或 THFT1、THFT2 等搭配是不行的。UPSHIFT 必须为 YES，除非 DBA 中的相应词全部大写。BLANK=N，表示当该字段（第二键）为空白时抑制主键（主词）而不必建立一个空白键。OCC=100 表示该重复字段至多可取 100 次。

这些索引语句在调试，修改词表过程中要反复使用，为避免反复手工键入，可对每一个索引在第一次键入后，对"DO YOU WISH TO SAVE SORT KEY DEFINITION?"答"Y"，然后对"PLEASE SPECIFY FILE NAME?"答文件名，以后做索引时，只要调出来用就可以了。如对主词，设保存索引语句用的文件名为 MTFILE，再对主词索引时，只要键入如下语句即可：

```
ISN=1/1000
OUTPUT=THMT
USE=MTFILE
```

（4）在 DATADICT 模块下，定义 KSAM 库 DBB。DBB 用于装入最后形成的，可供文献库调用的词表。DBB 库的头标定义中，库类型是 RD，文件类型是 THES，KSAM 文件名（KASM 中 FILE NAME）取任意 4 个字符，设为 DBBR，以后文献库调用词表的倒排文件根名应答此名。DBB 为库名。然后只要回答了键长和语种数目后，所需的全部字段系统就已自动定义了。在对字段操作的菜单中选择 LIST 或 SHOW，即可以看到系统自动定义的各个字段。我们沿用前例来进一步说明：回答键长为 40，语种数为 2，则系统就自动建立了下列字段：

字段名	下标	长度
MAIN TERM	X010	40
CN OF MAIN TERM	X020	6
CN OF LANGUAGE1	X030	6
CN OF LANGUAGE2	X040	6
ISN OF TERM	X050	6

X010 是这个 KSAM 文件的主键，供装入词表的具体内容。X020 是二级键。CN 为概念号，系统用它实现语种的翻译、查找和建立用代关系。

(5) 经过前述的步骤，就可以装词表了．沿用上例，做法如下．

在 THLOADER 模块下键入：

```
ENTER NAME OF THESAURUS KSAM DATA BASE OR EXIT—DBB
FILE THESAUR = THMT
FILE BT = THBT
FILE NT = THNT
FILE RT = THRT
FILE ANY = THANY
FILE FT = THFT
LANGUAGE EN = "E"，CH = "C"
END
```

其中第一个提示应答 KASM 库名；FILE 语句应给出的是对应的索引输出文件名，其中"FILE FT = "应给出第一种语言用代关系的索引输出文件名，而不是第二种语言的．LANGUAGE 语句中的"E""C"应与主词做索引时的后缀相一致．"EN""CH"在查询(QUERY）时用到，查询时用语句"LANG <语种列表>"来确定对哪种语言进行翻译．

键人前述内容后，系统就对 DBB 自动执行下述工作：将用以实现翻译功能的概念号分配给每个主词，标识出禁用词并装人相应使用词的概念号，建立 STRCKEYD 文件结构词代码和概念号．至此，就建立了多种语言的词表 DBB．系统在自动进行上述工作时，发现有不合法（词表结构不符合步骤（2）中的要求）的词，将显示其 ISN 号和相应的错误信息，用户应根据这些信息重新执行步骤（2）改错，再执行步骤（3)、(5)．

在上述过程中，对 DBB 库不能直接进行录人和修改．其中的数据内容是在 THLOADER 模块下装人的，因此要修改词表，增删内容，必须从步骤（2）开始修改 DBA 库，然后执行步骤（3)，对各个必要的字段做索引，再执行步骤（5)，用 THLOADER，重新装整个词表到 DBB 中．

(6) 在 DATADIC 模块下建文献数据库 DBC．DBC 中的词表控制字段，其性质应是：

VALIDATON（Y / N）—Y
ROOT NAME OF VALIDATION FILE—KSAM 文件名（例中为 DBBR）
OFFSET TO FIELD—4
LENGTH OF KEY TO EXTRACT—KSAM 文件主键的长度（例中为 40）
TYPE OF EXTRACTION（T / W / ）—选 TERM 或整个字段
TYPE OF INVERTED FILE（K / B）—K
ROOT NAME OF INVERTED FILE—KSAM 文件名（例中为 DBBR）
THESAURUS INVERTED FILE（Y / N）—Y

(7) 录人文献数据库的记录，其中词表控制字段的内容将验证文件（词表）的制约：必须用大写字符录人；只有词表中的主词能被接受；所有的非法词（包括禁用词）无法进人记录，成为字段内容．但非法词键人后，验证文件将其相邻的内容在屏幕上显示，供用户选择．验证文件中词的顺序是：按汉语拼音排序的中文主词，按字母顺序排序的英文主词，按汉语拼音排序的中文禁用词，按字母顺序排序的英文禁用词．前例中，若输人

"谷"，则屏幕显示"谷物"及顺序而下的 10 个主词供参照。若输人"CEREAL"，则屏幕显示"CEREALS"及以下的 10 个主词。

对文献库中词表控制字段的检索，可运用各种检索策略。MINISIS 具备了很强的多种语言词表功能：用一种语言的词查询，将命中多种语言关于同一内容的记录；可检索出包含某一主词的上位词、下位词、关系词的记录；用 ANY 词，可检索出相应 ANY 表中全部词的记录。如检索用词是禁用词，则会自动引导到使用词进行检索。用前例中的词标引，对文献库输人若干个记录。其检索效果如下：

进入 QUERY 模块，查询文献库中的词表控制字段（大小写均可）：

```
>=cereals
   CEREALS   P=2
   谷物   P=1
     1:     P=3       T=3

Q> bt cereals
   CEREALS   P=2
   谷物   P=1
   FOOD   P=1
     2:     P=3       T=3

Q> nt cereals
   CEREALS   P=2
   谷物   P=1
   RICE   P=1
   稻米   P=2
   WHEAT   P=1
   小麦   P=1
     3:     P=7        T=7

 Q> rt cereals
    CEREALS   P=2
    谷物   P=1
    AGRICULTURE     P=2
    HARVESTING     P=1
       4:     P=5       T=5

 Q> any agricultural production
    AGRICULTURAL PRODUCTION    P=1
    农产品    P=2
    CEREALS    P=2
   谷物    P=1
      5:      P=6       T=6

 Q> ft on
 Q> cereal
    CEREALS
    谷物    P=1
       6:     P=3       T=3
```

用中文检索，效果相仿，只是显示命中记录时，中文在前，英文在后。

在 QUERY 模块下，对文献库可用"DISPLAY [tag]@语句看词表结构内容，用"DISPLAY [tag]!@"语句看禁用词内容。

二、中英文双语种词表特例的讨论

MINISIS 有多种语言词表功能，中文也是被支持的语种之一，因此就可以很方便地用 MINISIS 来建立中英文词表。在计算机上建中英文词表之前，应有一本书面的中英文词表。目前，尚无国家标准的人口专业词表，因此我们以英文人口专业词表为蓝本，翻译整理成中英文人口专业词表。由于是双语词表，且中西文之间的差距远甚于两种西文之间的差距。碰到了诸如一词多译，多词一译等特殊情况，这里，把我们的处理方法提出与大家商榷。

(1) 如果对一个主词，英文与中文禁用词数目不一样，只要在两种语言禁用词录入时，重复次数不一样即可，其效果与中英文一一对应完全相同。例如：英文主词"BICYCLE"有一个禁用词"BIKE"，表示同一意义的中文主词"自行车"有两个禁用词"脚踏车"和"单车"，在对词表内容库录入时，按下面两条记录格式录入即可（沿用前面例子中定义的字段）：

INS=9
A010 MTE：bicycle
A020 MTC：自行车
ISN=10
A070 FTE：bike
A080 USEE：bicycle
A090 FTC：脚踏车
单车
A100 USEC：自行车

(2) 作为 (1) 的特例，会有这样的情况，一个英文主词（使用词）有一个或多个禁用词，但它们的中文译词相同，且只有一个。MINISIS 可用下述方法处理：禁用词是可重复字段，用代关系的记录允许只录入一种语言，这样做不会影响其以后在检索中的不同语言翻译功能。例：

INS=12
A010 MTE：aging
A020 MTC：老龄化

INS=13
A070 FTE：ageing
A080 USEE：aging

(3) 与 (2) 的情况相反，一个英文主词本身没有禁用词，译成中文却有一个主词和

多个禁用词，也可仿上例的方法处理．例:

INS = 14
A010 MTE : seal
A020 MTC : 图章
INS = 15
A090 FTC : 印鉴
　　印章
A100 USEC : 图章

上述这些问题的提出是由于中英文互译而引起的．翻译中，根据理想的词表规范化处理，英文词与中文词应一一对应．由语种不同引起的禁用词数不同则由禁用词字段是可重复字段来实现．MINISIS 对这些情况都能执行相应的翻译功能．

词表的建立和使用，大大加强了文献数据库的情报检索功能．我们把在工作中所做的、所想的，整理成本文，希望各位指正．

The Implementation of the SNOMED System using the MINISIS DBMS in Chinese-English

Wang Zhe
Ding Zhöian
Library
China-Japan Friendship Hospital
Computer Centre, ScienTech Information Centre
The Ministry of Machinery-Electronics Industry

Abstract: Systemized Nomenclature of Medicine (SNOMED) is one of the largest systems for classification and nomenclature of diseases in medicine today. It is significant to translate the name of its terms into Chinese and to implement it on the MINISIS DBMS using the English-Chinese version. This paper is divided into 3 parts: 1. Introduction of SNOMED, 2. Flowchart 3. System Design.

1. Introduction of SNOMED

Systemized Nomenclature of Medicine (SNOMED) is a coding system of nomenclature of medicine published by the College of American Pathologists and authorized by the College's Nomenclature and Classification of Disease Committee. It is one of the largest systems today, as compared with the International Classification of Disease (ICD) and the Chinese Classification of Disease (CCD), etc.

The SNOMED system has a database of nomenclatures used in medicine with codes for various clinical fields. There are 7 axes, including various terms and their codes. It is a large, useful, complete system which can be used in clinical research, medical records, for retrieval of filed clinical materials terms etc. Therefore, translating it into Chinese, to have a Chinese-English version of it, is of great significance and would strengthen international exchange.

We obtained the second edition of SNOMED data on computer tape delivered to our hospital through the kindness of the College of American Pathologists Computer Center. There are, in total, 44,581 records with the nomenclatures and codes in English.

Our translation of the system is accomplished by mounting the SNOMED database into a MINISIS database and adding the Chinese translations in a separate field. This way we created the first SNOMED DB in English and Chinese in our country. It might satisfy the need for this resource in our country. It is very convenient for communication of medical information between China and the outside world.

2. Flowchart

a. Analysis of the SNOMED tape format
b. SNOMED tape format converted to ISO 2709 format
c. Define the MINISIS RD and CD databases
d. Run ISOCONV
e. Run INDEX and INVERT for OFFLINE inverted files
f. Add a definition of the Chinese field, an online inverted field
g. Design a VPLUS worksheet for input
h. Input the Chinese field
i. SNOMED with Chinese field is converted to ISO 2709 format on tape for exchange
j. SNOMED with Chinese field is transferred to a micro-computer floppy disc for exchange

3. System Design

a. Condition for implementation of the system

In our China-Japan Friendship Hospital computer centre, we have an HP3000/70 equipped with an 8MB RAM DISC; 4 2000MB tape units; and 2 HP-VECTRA Chinese-English workstations. Online we use AdvanceLink and the Chinese-English version of MINISIS.

b. SNOMED tape structure:

Recording mode:	Fixed/Blocked
Record length:	128 bytes
Blocking factor:	8 (1024 bytes)
Tape identification:	unlabelled
Tape format:	ASCII
Tape byte count:	5.7 MB

c. Record content structure

Tape Mark	First Character	Number of Records
Copyright notice	binary 00000001	6
Disease axis	D	6000
Etiology axis	E	5785
Function axis	F	7948
Occupation axis	J	1507
Morphology axis	M	8399
Procedure axis	P	5096
Topography axis	T	9846
Total byte count		44,587

4. Details of system implementation

a. Analysis of SNOMED tape format

D0001001 Disease type of Category unknown

D0002001 Disease type of Category not application

D0003001 Disease type of Category not assigned

b. Program Design

Language: FORTRAN

Program Flowchart:

DATA

READ RECORD

COMPUTE FIRST FIELD LENGTH FL1
FIELD1 = RECORD[0:FL1],FL1 = FL1 + 1
COMPUTE SECOND FIELD LENGTH FL2
FIELD2 = RECORD[FL1:FL2]

DIR1[0:2] = "010",DIR[3:6] = STR("48",4)
DIR1[7:11] = STR((FL1-1),5)
DIR2[0:2] = "020",DIR[3:6] = STR(("48" + FL1),4)
DIR2[7:11] = DTR(FL2,5)

REC = FL1 + FL2 + L + LOD + 2
HEADER[0:4] = STR(RECL,4)

RE-RECORD:ISO FORMAT RECORD,WRITE DISC

c. Converted ISO 2709 tape format

00070NA > 1100049 > 45 > 100000900000200001300009/D0387002/Shipyard eye
00070NA > 1100049 > 45 > 100000900000200002100009/Epidemic pleurodynia
00070NA > 1100049 > 45 > 100000900000200001700009/Bornholm disease

d. Define the RD and CD data bases in MINISIS

```
              LISTDDT DATA DEFINITION
                                        L  P  I  I  I  N  T  L  I
                                        E  R  N  N  N  U  Y  E  N
                                        N  O  V  V  V  M  P  N  V
                                        G  H
    FIELD NAME     NEMONIC    TAG       T  P  O  O  T  E  E  E  F
                                              N  F  Y  X  X  X  I
                                                 F  P  T  T  T  L
                                                    E  R  R  R  E
--------------------------------------------------------------------------
    Data Base : wz

    CODE           CO     A100        8                    Y  B  10    8 IDW1
    ENOTE          EN     A200      120                    Y  B  10   50 IDW2
    CHINESE NAME   CH     A300      100  Y  Y                 B  10   40 IDW3
--------------------------------------------------------------------------
                     CD DATABASE HEAD

    LENGTH OF RECORD SEPARATOR  LENGTH OF FIELD SEPARATOR  UNBLOCKED TAPE (Y/N)
           1                          1                         Y
--------------------------------------------------------------------------
                       CD FIELD TAG

    MINISIS TAG  ISO TAG  LENGTH OF ISO FIELD   LENGTH OF INDICATOR   "FROM"
       A100      100             8                     0                0
       A200      200            120                    0                0
       A300      300            100                    0                0
--------------------------------------------------------------------------
```

e. Run ISOCONV to create MINISIS DB

```
:FILE T;DEV = TAPE;REC = -200,32,F,ASCII
DUMP FROM = CD, TO = SNOMED
Enter command to select records
SUBSET = all
***Please Mount ISO tape***
Total records processed = 44,587 Write = 44,587 Rejected = 0
Time = 20 minutes
```

f. Run INDEX and INVERT for OFFLINE inverted files

INDEX:

```
DB NAME: SNOMED
RECORDS SELECTED: ISN = 1/50000
OUTPUT = OUTF,SORT = YES
KEY = CODE,LENGTH = 8,UPSHIFT = YES
FIELD = A100
END
***
```

INVERT:

```
INDEX/TLOG FILE NAME: OUTF
INDEX/TLOG FILE NAME: OUTF1
KEY FIELD = IDW1
KEY FIELD = IDW2
NUMBER OF INDEX RECORDS = 44,587
NUMBER OF KEYS WRITTEN = 44,587
```

g. Design a VPLUS worksheet for input

```
        :Run formspec.pub.sys
----------------------------------------------------------------------
        ISN [     ] code [        ]
        ENOTE [                                                      ]
        CNOTE [                                                      ]

        F4 : TO SEE THE FUNCTION
        F5 : TO BREAK THE PROCESS OF UPDATE
        F8 : TO ENTRY FIELD TO THE DATABASE

---------------------------------------------------------------------
---------------------------------------------------------------------
NUM[     ] LEN [      ] NAME [          ] ENH [HI ] FTYPE [  ] DTYPE [     ]
INITIAL VALUE  [                                                         ]
        *** PROCESSING SPECIFICATIONS ***
---------------------------------------------------------------------

    ISN,CODE,ENOTE (English name of disease): FTYPE[D] (Display)


          CNOTE (Chinese name of disease): FTYPE[O] (Option)
    Name of parameter file-parmfile : *form=wzform
```

h. Input the Chinese field

During the input of the Chinese field, after entering 200 records, a print copy was made to use in checking for errors. Error correction was then undertaken and a back-up was done to prevent data loss.

Because of the professional feature of medicine, there were not always Chinese words in the the Chinese Character bank to match the English term. We made corresponding characters to solve this problem.

i. SNOMED with Chinese field is converted to ISO 2709 format on tape for exchange

```
    :FILE ISO;DEV = TAPE
```

Enter module 10 (ISOCONV)

```
    LOAD FROM = SNOMED, TO = CD
    SUBSET = all
    ***PLEASE MOUNT ISO TAPE***
    TOTAL RECORDS PROCESSED = 44,587 WRITTEN = 44,587
    REJECTED = 0
    TIME = 30 MINUTES
```

Conclusions

Completion of this subject, familiarizes ourselves with MINISIS relational DB G version 7 main function model. We defined the two types of data bases, the RD and the CD, using DATADICT. We practiced the English-Chinese data entry and modification in UPDATE and we designed an entry worksheet using the VPLUS system. There are now 44,581 records in the data base. Using ISOCONV, we loaded the ISO tape of SNOMED data into MINISIS and used the DUMP command to convert from the CD to the RD. We then defined and entered the 44,581 translated terms into the Chinese term field completing SNOMED. We also used LOAD to go from the RD to the CD and then on to the ISO tape format to promote the dissemination of this work. We have become familiar with the various functions of MINISIS through our work. Though continuing to explore and practice, we have gained some experience leading us to conclude that the MINISIS functions are convenient and practical. Unfortunately, the time it takes to enter the Chinese fields is long, we found we could enter about 50-60 records an hour. That is about 10 Chinese words or 40 Chinese characters an hour on average.

Aknowledgement: We especially express our great gratitude to Dr. Li Ensheng, Director of the Department of Medical Records of the China-Japan Friendship Hospital for his honest support of us. We also thank Mr. Jiang Xiangdong of the MINISIS Resource Centre. Thanks also to Director Mai Zu Yin and Vice Director Li Shiliang of the China-Japan Friendship Hospital Computer Centre.

References:

MINISIS DBMS Volumes 1, 2, 3
MINISIS G Version

以MINISIS 实现系统医学命名法 (SNOMED) 中英文对译系统

王喆　丁志坚，　中日友好医院图书馆,机电部情报所计算中心

摘要:系统医学命名法是世界现用医学疾病分类较大系统之一,是临床疾病分类编目不可缺少的词表。将其汉化,建成英汉对照词表起到中英互译功能具有深远意义,填补了我国这项工作的空白。整个过程是在HP-3000/70机上,在MINISIS功能的基础上实现的,方便实用。本文分三部分:一.SNOMED 介绍 二.流程图 三.系统实现

一. SNOMED 介绍

系统医学命名法(Systematized Nomenclature of Medicine;SNOMED)是由美国疾病系统和分类委员会授权美国病理学者学院编排出版的医学疾病名称编码系统。与其他系统,如,国际疾病分类(简称ICD),中国疾病分类(简称CCD)等均是现代较大分类系统之一。

SNOMED 系统是实用于医学科研、疾病编目、检索等多方面使用的编码系统化的名称库,是具有七个轴心包括临床使用各个领域的名称和编码。是一个比较庞大的适于计算机名、码查询检索使用的较完整的系统。因此,将其汉化,对临床、科研、病案管理、疾病编目等工作有极大参考意义,并可加强国际间交流。

数据来源是由美国病理学者学院计算中心研制的向我院病案室提供的SNOMED磁带文件,共存44,581条英文名称及编码。我们的中英文对译系统是用SNOMED第二版磁带建库后再增加中文翻译字段,建成我国第一部具有英汉两种文字的SNOMED数据库,它填补了我国这项工作的空白。

二. 流程

1. 分析SNOMED 非标准磁带格式
2. 程序设计,将非标准磁带格式转换成ISO2709格式
3. 进入MINISIS管理系统.定义:RD,CD数据库
4. 运行ISOCOVN处程序,完成数据装库
5. 做脱机倒排文件
6. 在DATADDICT中,增加定义一个汉字字段,为联机倒排文件
7. 设计VPLUS输入格式文件
8. 录入汉字字段,最终完成SNOMED中英翻译数据库
9. 把SNOMED中英翻译数据库转换成ISO2709磁带文件,以便国际交换
10. 把SNOMED中英翻译数据库转入微机软盘,以利推广使用

三. 系统实现

1. 系统实现的环境

我中日友好医院计算中心是HP-3000/70机,内存8MB,外存DISK 4 台 2000MB,2台磁带机,HP-VECTRA 汉英工作站,联机软件ADVLINK。IDRC 无偿提供的MINISIS中西文文献数据库管理系统。

2. 原始数据磁带结构

计录方式	FIXED/BLOCKED
计录长度	128 BYTES
块因子	8 (1024 BYTES)
磁带标识	UNLABELLED
磁带代码	ASCII
总字节数	5.7 MB
磁带规格	1/2 INCH,1600BPI,9 TRACK MODE

3. 计录内容结构

SNOMED 是有七个轴心的庞大的包括解剖(T)，病因(E)，机能(F)，形态(M)，疾病(D)，操作(P)，及职业(O)等各有八位码的名称库。共计有44,581个词条.如下所示：

磁带计录内容	第一字节	计录数
版权标记	00000001	6
疾病轴	D	6000
病因轴	E	5785
机能轴	F	7948
职业轴	J	1507
处置轴	P	5096
形态学轴	M	8399
解剖部位轴	T	9846

4. 系统实现的详细过程

(1) 分析SNOMED非标格式磁带

```
定义磁带文件与磁盘文件:
FILE T;DEV=TAPE;REC=-128,8,F,ASCII
BUILD F1;REC=-128,8,F,ASCII;DISC=50000
由磁带将数据拷贝到磁盘:
:FCOPY FROM=*T;TO=F1
由DISC读到终端:
:FCOPY FROM=F1;TO
打印机打出数据
:FILE L;DEV=LP
:FCOPY FROM=F1;TO=*L
```

非标准磁带格式

```
-----------------------------------------------------------
D0001001    Disease type of category unknown
D0002001    Disease type or category not application
D0003001    Disease type or category not assigned
   :                              :
-----------------------------------------------------------
```

(2) 程序设计，将非标准格式磁带转换成ISO2709格式

程序设计语言:FORTRAN
框图如下:

```
              -------------
               原数据
              ----------
               读记录
              ----------
-----------------------------------------------
计算第一个字段长度 FL1
FIELD1=RECORD[0:FL1],FL1=FL1+1
计算第二个字段长度 FL2
FIELD2=RECORD[FL1:FL2]
-----------------------------------

DIR1[0:2]="010",DIR[3:6 ]=STR("48",4)
DIR1[7:11]=STR((FL1-1),5)
DIR2[0:2]="020",DIR[3:6]=STR(("48"+FL1),4)
DIR2[7:11]=STR(FL2,5)
-----------------------------------------------
RE-RECORD:ISO 格式化原磁带，写磁盘
-----------------------------------------------
```

（3） 转化后的数据格式如图:(ISO2709格式)

```
-------------------------------------------------------------------------------
00083NA>>>1100049>>>45>>1000009000002000002600009/D0386002/Echo virus infection
00089NA>>>1100049>>>45>>1000009000002000003200009/D0386004/Enterovirus echo,
00070NA>>>1100049>>>45>>1000009000002000001300009/D0387002/Shipyard eye
                :                           :                        :          :
-------------------------------------------------------------------------------
    头标区    24位       目录区1    12位        目录区2    12位        数据区
-------------------------------------------------------------------------------
1-5    记录长度            1-3   标识符长度      1-3   标识符长度      1-8   代码段
6      记录状态            4-7   字段长度        4-7   字段长度        9-     疾病名称字段
7      系统代码            8-12  字段相对位置    8-12  字段相对位置
8-10   未用区                    数据起始地址          数据起始地址
11     指示符
12     标识符
13-17数据区起址
18-20未用区
21     目录区长度的长度
22     目录长度
-------------------------------------------------------------------------------
```

（4） 送入MINISIS管理系统定义RD,CD数据库

```
          LISTDDT DATA DEFINITION
                                      L  P  I  I  I  N  T  L  I
                                      E  R  N  N  N  U  Y  E  N
                                      N  O  V  V  V  M  P  N  V
                                      G  M
                                      T  T  O  O  T  E  E  E  F
                                               N  F  Y  X  X  X  I
FIELD NAME     NEMNIC     TAG                     F  P  T  T  T  L
                                                     E  R  R  R  E
-------------------------------------------------------------------------------
DATABASE : WZ

CODE             CO         A100   8           Y  B  10    8   IDW1
ENOTE            EN         A200   120         Y  B  10 W  50  IDW2
CHINESE NAME     CN         A300   100 Y Y        B  10    40  IDW3
-------------------------------------------------------------------------------
                         CD DATABASE HEAD

LENGTH OF RECORD SEPARATOR    LENGTH OF FIELD SEPARATOR   UNBLOCKED TAPE (Y/N)
            1                              1                        Y
-------------------------------------------------------------------------------
                         CD FIELD TAG

  A100            100             8                        0              0
  A200            200             120                      0              0
  A300            300             100                      0              0
-------------------------------------------------------------------------------
```

（5） 运行ISOCOVN处理程序,完成数据装库

```
:FILE T;DEV=TAPE;REC=-200,32,F,ASCII
>DUMP FROM=CD,TO=SNOMED
ENTER COMMAND TO SELECT RECORDS
>SUBSET=ALL
*** PLEASE MOUNT ISO TAPE ***
TOTAL RECORD PROCESSED=44,587 WRITE=44,587  REJECTED=0
约 20 分钟完成装库
```

(6) 做联机倒排文件

a. 在MINISIS 菜单中选索引功能模块3:

```
DB NAME : WZ;ISN=1/50000
OUTPUT=OUTF,SORTING=YES
KEY=CODE,LENGTH=8,UPSHIFT=YES
FIELD=A100(代码字段)
END
***

OUTPUT=OUTF1,SORTING=YES
KEY=ED,LENGTH=120,UPSHIFT=YES
FIELD=A200,TYPE=W,STRIP=N (英文字段)
END
***
```

b. 在MINISIS菜单中选倒排模块18:

```
选倒排功能:I
INDEX/TLOG FILENAME:OUTF
键字段:IDW1 代码段倒排文件根名
       IDW2 英文字段倒排文件根名
NUMBER OF INDEX RECORDS=44,587
NUMBER OF KEYS WRITTEN=44,587
```

(7) 在MINISIS定义模块中增加定义一个汉字字段，为联机倒排，其参数如前述。

(8) 设计VPLUS输入格式文件

```
:RUN FORMSPEC.PUB.SYS
--------------------------------------------------------------------
 ISN  [          ]       CODE  [                     ]

 ENOTE  [
 CNOTE  [                                                          ]
                                                                   ]
 注     F4 : TO SEE THE FUNCTION
        F5 : TO BREAK THE PROCESS OF UPDATE
        F8 : TO ENTRY FIELD TO THE DATABASE
---------------------------------------------------------------------
NUM [   ] LEN [  ] NAME [       ] ENH [HI]   FTYPE [  ]  DTYPE [    ]
INITIAL VALUE  [
    *** PROCESSING SPECIFICATIONS ***                               ]
---------------------------------------------------------------------
ISN,CODE,ENOTE(英文字段名): FTYPE [D](Display)
         CNOTE(中文字段名): FTYPE [O](Option)
参数方程为:*FORM=WZFORM
```

(9) 录入汉字字段，最终完成SNOMED中英翻译数据库

在汉字字段录入过程中，每录入200条记录后，打印出拷贝进行校对，修改错误 。
每天做数据文件备份，以防磁盘损坏，录入数据丢失。
在录入汉字字段时，因医学专业名词多，出现汉字字库没有的汉字。我们用汉英工作站所具有的造字功能，解决了问题。

(10) 把建好的中英对译MINISIS库转换成ISO2709磁带格式，以进行国际交换。参数如下：选模块 20，定义CD数据库

```
      ------------------------------------------------------------------------
                             CD DATABASE HEAD
      LENGTH OF BLOCK LENGTH FIELD : 0        UNBLCKED TAPE (Y/N) :N
------------------------------------------------------------------------------
                             CD FIELD TAG
MINISIS TAG    ISO TAG   LENGTH OF ISO FIELD    LENGTH OF INDICATOR    "TO" FIELD
 A100           100            8                        0                  0
 A200           200            120                      0                  0
 A300           300            100                      0                  0
------------------------------------------------------------------------------

      :FILE ISO;DEV=TAPE
      进入模块 10 (ISOCOVN) :
      >LOAD FROM=SNOMED,TO=CD
      >SUBSET=ALL
      *** PLEASE MOUNT ISO TAPE ***
      TOTAL RECORDS PROCESSED=44,587 WRITTEN=44,587 REJECTED=0
      约30分完成转换磁带。
```

结论

本课题的完成，使我们实践、熟悉了MINISIS关系型数据库管理系统G版本的七个主要功能模块。在DATADICT 定义了RD，CD两种类型数据库参数。在 UPDATA 进行了两种文字，英汉字段的录入，修改，并设计了VPLUS录入格式进行汉字字段的录入。在QUERY中，在44,581条记录库中查询记录，从三个字段中（中、英、代码）任何一个字段都可查到所需的记录。我们既实践了联机倒排字段功能（汉字字段是联机倒排）也实践了44,581条记录的脱机倒排功能（代码与英文字段是脱机倒排字段）。ISOCOVN中，我们首先把ISO磁带中的SNOMED数据装入MINISIS库中，使用DUMP命令实现由CD到RD的转换功能。然后，定义、录入了44,581条汉字翻译字段，最终形成了具有英汉对照的SNOMED词表数据库。为了便于推广使用，我们又以LOAD命令实现由RD经CD到ISO磁带的数据转换功能。这些过程的实践，使我们熟悉了MINISIS的各种功能．不断地探索、实践，使我们积累了不少经验。同时，我们感到MINISIS功能全面，方便，好用。在汉字字段的录入过程中，从字段的录入到再次出现下一个录入时，这个间隔有等待时间较长的感觉。汉字的录入时间是50-60条/小时左右（每条记录汉字字数约10个）

在此，中日友好医院病案室主任李思生教授，计算中心主任袁福荫，付主任李士良热情支持，MINISIS中心江向东老师予以指导，特致谢意。

参考文献：MINISIS F.VOLUME I,II,III
MINISIS G VERSION

COMPUTERISATION AT THE NATIONAL LIBRARY OF MALAYSIA: FUNCTIONAL SPECIFICATIONS FOR AN INTEGRATED LIBRARY SYSTEM

Lim Chiew Aun
National Library of Malaysia

1. BACKGROUND

Computerisation activities began at the National Library of Malaysia in 1978 with the processing of its cataloguing data using MALMARC (Malaysian Machine Readable Catalogue). Subsequently, the MALMARC computer facilities were also used for the production of the Malaysian National Bibliography and the creation of a Union List of Serials.

Following this, microcomputers of different makes were acquired at different times and used for various purposes such as the production of the Malaysian Newspaper Index, input of cataloguing data on diskette for the Malaysian National Bibliography, the production of reports, leaflets, etc, and for accessing external databases. It was in 1985 that MINISIS was used for the production of the Malaysian Periodicals Index and later the maintenance of the databases for the Standard List of Malaysian Government Bodies and Standard List of Malaysian Statutory Bodies.

Hence, the scenario was that of isolated computer systems maintaining different files and databases. As these systems are not linked, accessing them would mean the need to move from one system to another.

2. The Need for an Integrated Library System

The library started looking seriously into acquiring a total library management system for the management and control of its functional activities when development funds became available for its computerisation programme under the Fifth Malaysian Economic Plan (1986-1990). Studies and recommendations were then made by the Unesco consultants' Michael Sherwood (1984) and Kerry Webb (1986) specially for the computerisation of the National Library. Following these studies, a special task force was set up within the library to make a detailed study of the functional requirements of its various activities, processes, and services. The studies (See Appendix I - Functional chart of National Library of Malaysia) undertaken pointed to the need for a single central database with multiple interactive access and this need can be met by an integrated library system.

Under the manually operated system, the major problem faced by the Library was that of data duplication related to activities such as duplicate checking, maintaining statistics, record updating, etc, Each user keeps different versions of those data and has few relationship with other users. For example, in the activities of acquisition, cataloguing, and indexing, there are 27 files of sets and subsets of essentially the same data. The speed of data flow through such a system is very slow and is conditioned by this constant duplication. Hence, there is a dire need for the library to acquire a library management system which will enable it to effectively and efficiently manage and coordinate all its functional activities and information system. Such a system should minimise data duplication and allow the application of transactions to different functional activities with no delay.

3. Functional vs. Data Integration

Research on the literature indicates that there are two categories of integrated library systems, namely data and functional integration. first category of integrated systems is modular in nature. In such a system, functions such circulation, cataloguing, serials management, OPAC, etc. are designed as separate modules each using a common, non-redundant, multi-purpose database system. However, the second is a higher category of integration which does not have modules but have subsystems.

The key feature of a functionally integrated system is that it is not necessary to exit one subsystem prior to executing the functions that normally reside in another subsystem. For example, a user logged on to the cataloguing subsystem of a functionally integrated system can easily execute functions residing in the circulation function without leaving the cataloguing subsystem. Such a system is attractive and time-saving as it provides a noncaptive man/machine dialogue which allow users, if they choose, to ignore the current dialogue, and initiate new dialogues at any time. As the National Library's computer terminals are expected to be scattered throughout several floors in its new building the flexibility of a functionally integrated system would certainly be a great advantage as it allows transactions on multiple functions to be executed at scattered service points.

4. Functional Specifications: Integrated Library System

After all the research and studies, a proposal was made to acquire an integrated library system, and approval for this was given by the relevant authorities at the end of 1987. As required by Government Regulations, a tender exercise was conducted in early 1988 to acquire the proposed system. Among the tasks to be undertaken in such an exercise was that of preparing a detailed specification to meet the functional requirements of the library.

A technical committee was set up and commissioned to prepare this technical specification which make up the core content of the tender document. Several sub-committees were also formed to study and spell out the detailed specifications necessary for each of the functions to be automated.

The final document on the technical specification consists of 4 parts namely, Part I which provides background information on the National Library's functions, collection, processing needs, etc.; Part II which provides the general specifications on the scope and general requirements for hardware, software, system developments, and customisation work required etc.; Part III covers the "Mandatory Requirements" while Part IV covers the "Essential and Desirable features". These two parts provide the core information on the functional specifications to be fulfilled by the integrated library system. A diagrammatic representation of the total integration of all the functions as required by the National Library is given at Appendix II. For each function the major features looked for are described under their respective headings.

4.1 Mandatory Requirements

These are standard functions which the integrated library system must have to support the major functional activities of the library. They are basically the following:

4.1.1 Acquisition
4.1.2 Cataloguing;
4.1.3 Serials management;
4.1.4 Online Public Access Catalogue (OPAC);
4.1.5 Interlibrary loans (ILL);
4.1.6 Circulation.

Generally, the system must cater for a single database with total functional integration to preclude redundant data storage as well as permit incorporation of additional features with the same degree of integration. It must allow all transactions to be initiated at any screen display. MARC formats with both fixed and variable length fields are to be fully supported, and the serials system must provide for the new MARC format for serial holdings and locations. Besides printed lists and bibliographies to be produced according to specific formats, it must also be able to produce printouts for the publication of the National Bibliography of Malaysia.

The system s also required to handle the following:

(i) Online processing and online enquiries;

(ii) Interactive multi-user, multitasking operations in a time-sharing environment;

(iii) Allow bibliographic records to be merged such that linkage to item records are automatically changed (retrospective conversion);

(iv) Allow records of all types to be fully modifiable at the character or data element level, and to be deleted from data base when no longer needed, without any programme intervention.

4.1.1 Acquisitions

The acquisitions function is required for the recording and control of the acquisition of library materials via purchase, gift and exchange, institutional repository arrangements, and legal deposit. Core functions of creating an order, allocating suppliers, batching and printing orders, chasing overdues are all expected as standard features. It should include the maintenance and updating of suppliers' and donors' files; fund accounting and budget files; updating of acquisition records; bibliographical searching and duplicate checking facilities; recording of part delivery and part payment of orders; currency conversion; generation of forms, letters, and statistical reports, at specified intervals and according to specified formats.

On the financial aspect, the system is expected to do basic arithmetical calculations such as totalling of order price from item price and number of copies, handle discounts and taxes, as well as maintain details of amounts committed, spent, free and committed balances. The library should be able to search and use the existing catalogue data to create an order record; and conversely, once an item is received, the order data, particularly the bibliographical fields, can be copied back and used as the basis of the catalogue record.

4.1.2 Cataloguing

Features looked for under the cataloguing function include the creation, amendment, enquiry, deletion, maintenance and control of bibliographic records in one integrated data base. The system must fully support MARC formats as well as allow the MARC bibliographic and authority records from the National Library's database in MALMARC to merge with the Malaysian National Bibliography database, as one database. Authority files and bibliographic files are to be linked under the authority control facility so as to provide global update of authority headings when modifications are made. Various levels of display of the bibliographical description is to be made available for on screen viewing and printouts. All records should be retrievable by Boolean, free text, and truncated searches.

4.1.3 Serials Management

Basically it is required to provide serials control and subscriptions management with the logging of receipt of each issue of a serial as it arrives, as well as offer full prediction capabilities which could be based on frequency codes and prediction patterns set up for each publication. It should recognise and facilitate the creation and maintenance of bibliographic records and management of serials acquired through purchase (subscription), gift and exchange, institutional repository arrangements, and legal deposit. Authority control and thesaurus creation should be catered for and the bibliographic records should conform with international standards such as MARC and ISBD for serials.

It should reflect the library's total serial holdings and the records created should provide core information to other functions like OPAC, Circulation, Indexing, etc. Fund accounting and budget control facilities are expected, as well as the generation of forms, letter, statistical reports and other management information reports as specified by the library.

4.1.4 Online Public Access Catalogue (OPAC)

The function has to be user-friendly and give specific information about the library's holdings. Basic instructions for accessing the data base should be short and simple to enable inexperienced users to perform enquiries with minimum assistance and in the minimum time. It should basically enable the user to trace an item by author, title, subject and series, and to ascertain its availability and location. HELP facilities are to be available at any stage of a search strategy. Variant spellings and display of authority files to facilitate searching should be handled by the system. Facilities to save search sets and answers to selected previous enquiries should also be provided.

4.1.5 Interlibrary Loans

Full administration of ILL is to be provided so that requests can be tracked from the time they are requested by the borrower to their ultimate dispatch back to the library of origin. The library should be able to fulfil interlibrary loans, photocopy or micro form requests by electronically transmitting through the system, or by the printing and despatch of interlibrary loan forms. Alerts and visual messages should be signalled for reserved titles as well as overdue payments and loans. There should also be automatic calculation of due date and payment chargeable. The system is required to generate statistical reports as well as printouts of forms and letters for overdue loans, recall notices, payment notices, interlibrary loan forms, etc.

4.1.6 Circulation

Basically, this module is required to provide control of all items in stock according to the loan policies of the library. Facilities provided should include charging (issues) and discharging (returns) of loans; renewals; patron registration information; circulation statistics/reports; reservations and recalls; write off/lost items/binding/stocktaking data; fines; deposit accounts; and bibliographic services and patron enquiry (OPAC). The library should be able to set up loan periods, borrower entitlement, fines due, overdue routines and contents of notices, and other parameters, in compliance with its policies for different loan and borrower categories. Automatic calculation of due dates, fines, etc. in relation to "dead days" of the calendar (holidays, library closed, etc.) should be provided.

The system must provide instantaneous checks and alert systems on items and borrowers on issue and return at different service points, and must be able to put items and borrowers on hold for various reasons such as reservations, overdue returns, unpaid fines, suspended loan privilege, etc. However, authorised staff should be allowed to override holds, and renewal policies, when necessary, and with the appropriate security provisions. It should be able to generate management information and statistical data on the utilisation of the collection, collection of fines, reservations received, as well as printouts of overdue and recall notices, receipts for payments received in part or in full, etc. Barcode input should be supported as well as keying in of data with provision for a back-up system in the event of system breakdown, or power failure.

4.2 Essential and Desirable Features

In addition to the major functions outlined above, specifications are also proposed for essential and desirable features which encompasses functions that require some customisation and are important, but are not mandatory to the system. These include the control and management of the following:

4.2.1 International Standard Book Numbering (ISBN) System, and International Serials Data System (ISDS);

4.2.2 Manuscripts management;

4.2.3 Indexing module;

4.2.4 Selective Dissemination of Information (SDI);

4.2.5 Conservation.

The major features looked for in these functions are described separately under their respective headings.

Generally, the system is expected to have the capacity and capability to integrate bibliographic data presently generated by the library's indexing activities, namely the Malaysian Periodicals Index, Index to Malaysian Conferences, and Malaysian Newspaper Indexes, or alternatively to support MINISIS software. It should be able to process bibliographic data in non-Roman scripts such as Jawi, Arabic, Chinese and Tamil, as well as provide HELP screens in the local languages. The system must be parametized to allow incorporation of local library policy into the system performance, and parameters must be easily changed without programmer intervention. The library integrated system should also be able to run concurrently with other applications software used by the library e.g. Data Base Management System, wordprocessors, statistical packages, office automation, etc.

Some customization of software will be required to address the special needs of the National Library such as processing of manuscripts, conservation, special indexing services, maintaining reference enquiry records, maintaining Selective Dissemination of Information (SDI) Services, maintaining the pamphlets collection, the United Nations collection or producing various reference sources and directories. Customization will also be required to produce the publication formats of the Malaysian Periodicals Index, Malaysian Newspaper Index, and Index to Malaysian Conferences.

4.2.1 ISBN and ISDS

Inventory control for used and unused ISBN and ISDS numbers and the validation of these numbers are required, as well as file maintenance with facilities to create, add, delete, and store data. Search facilities are to be provided to check and verify the status of items to be assigned ISBN/ISSN numbers. It should also be able to generate statistical reports and listings of publishers and ISBN/ISSN numbers assigned by language or for specified periods.

4.2.2 Manuscripts management

Basically the system is required to create and maintain a single data bank on Malay Manuscripts regardless of their location as well as create and main-cain a documentation file of the National Library's Malay manuscript collection. It should facilitate the recording and control of manuscripts acquired through purchase, gift, or trusts/bequests, and the maintenance of budget files. Multiple access points are to be provided for easy information retrieval of records, as well as authority control with facilities for amendment of any authority heading to be reflected globally in the bibliographic records.

An added feature required for a manuscript file is the ability of the system to accept both Arabic and Jawi Scripts as well as allow the intermixing of these scripts with Roman Scripts. The system should also be able to produce the National Library's Catalogue of manuscripts as well as a union catalogue of all manuscripts documented. Various statistical data and reports are to be produced for management information.

4.2.3 Indexing

This function would provide for the creation, addition, searching, amendment, deletion, maintenance and bibliographical control of articles in Malaysian serials such periodicals, journals, newspapers, and working papers of conferences, seminars, and workshops. It should be able to integrate existing records of the Malaysian Newspaper Index, Malaysian Periodicals Index, and the Index to Malaysian Conferences, as well as provide for the future enhancement of facilities and features to establish a national indexing system through networking. Boolean and free text searches are to be provided to facilitate information retrieval and keep track of articles indexed and their location in the library. Statistical data and various lists and reports are to be generated either for publication or for management information.

4.2.4 Selective Dissemination of Information (SDI)

The system is required to create and store "search profiles" and to run these on demand against new entries to the database. Bibliographical lists are generated and mailed to the respective user to keep him up-to-date and well-informed on new material in his subject or interest area. An estimation of the number of records matching a search should also be included in this function to assist in decisions for refining search strategies.

4.2.5 Conservation

This function is intended to provide online information on the conservation, binding, and microfilming of library materials in one integrated bibliographical database. There is to be built into the system some form of automatic identification of volumes or items ready for conservation, binding, and microfilming, based on the bibliographic descriptions input into the system. Control and maintenance of records (job sheets) on the flow of materials through the various processing activities are to be provided. Instructions and specifications for conservation, binding, and microfilming are to be included. Various statistics and reports are also to be provided by the system.

5. Conclusion

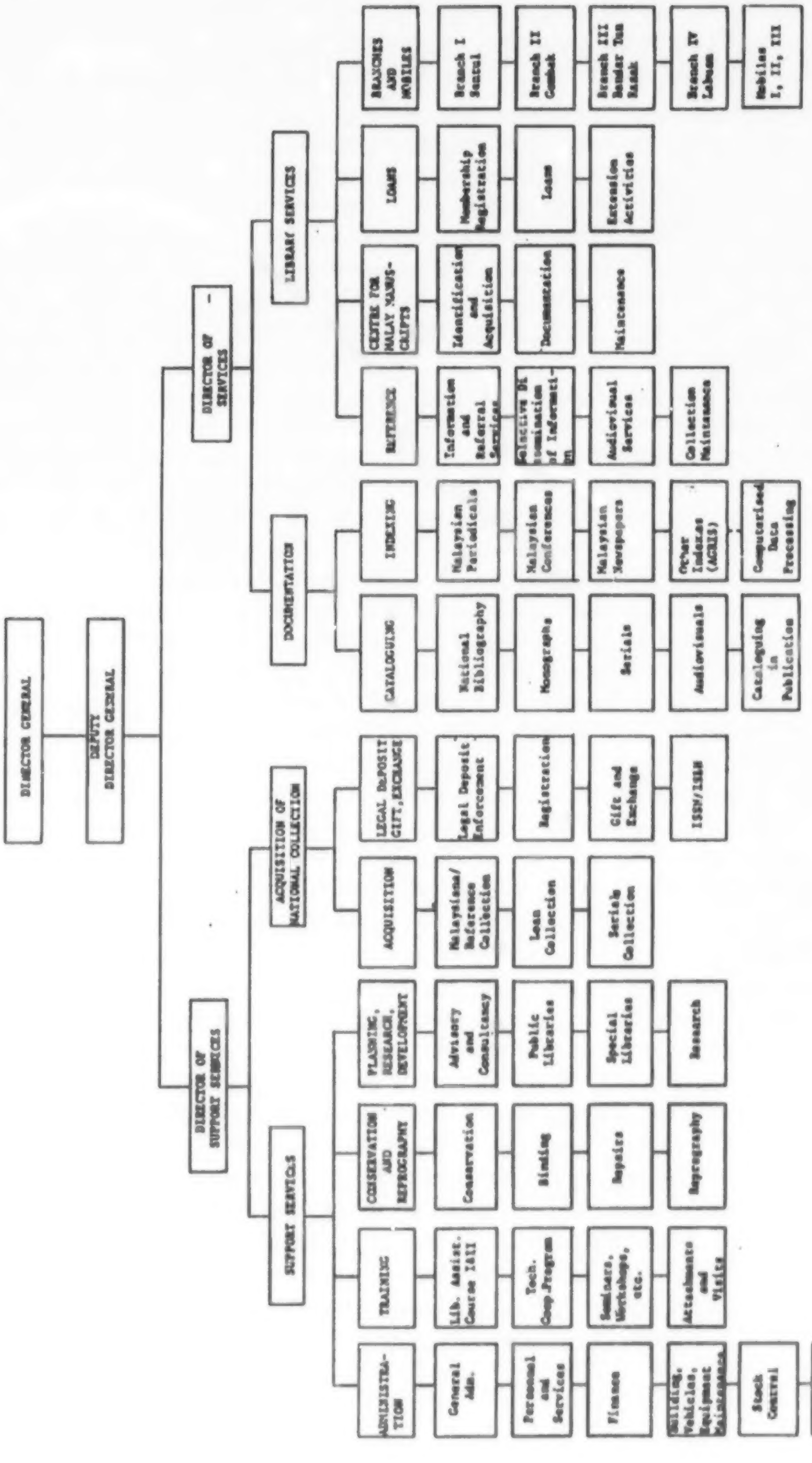
Functional specifications for an integrated library system have to be detailed in order to facilitate the evaluation and selection of a suitable system that can meet the functional requirements of the library. A functionally integrated system would allow single data entry with multiple access and provide different views of the data without the need to exit from one subsystem or function to another. This would reduce work duplication and increase productivity as well as economise on access time. The basic functions like Acquisitions, Cataloguing, Serials Management, Online Public Access Catalogue, Interlibrary Loans, and Circulation, are to be supported and integrated within the system. Some of the other functions such as the ISBN & ISBD systems, Manuscript management, Indexing, SDI, and conservation are functions specifically related to services and activities conducted by the National Library and would need to be customised for its use. For indexing, there is the option given to link the new library management system currently in use for this function.

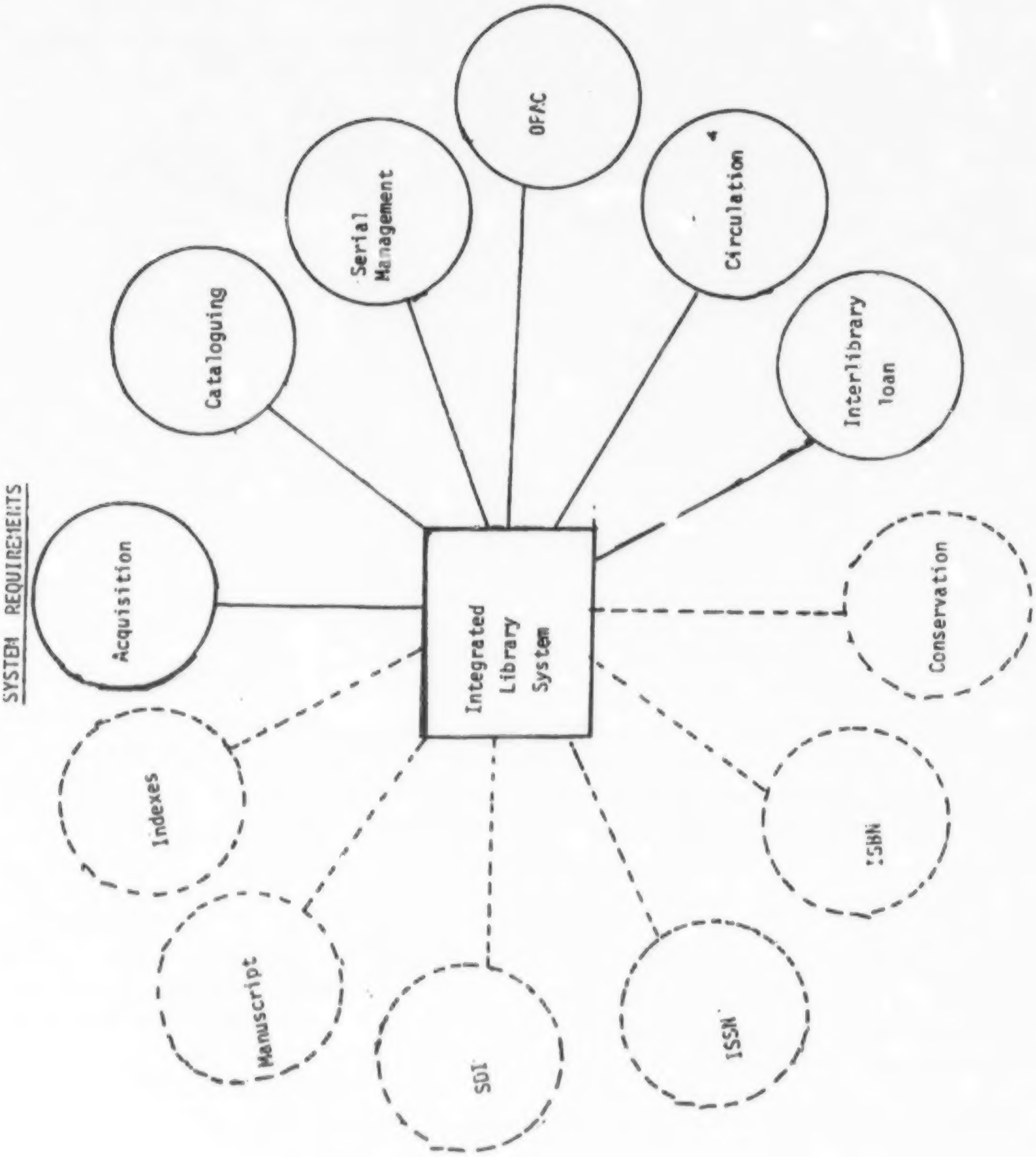
Based on the technical specifications, the National Library made evaluations of several integrated Library Systems before finally deciding to acquire the Virginia Tech Library System (VTLS), a functionally integrated system, operating on a HP3OOO Series 950 minicomputer. In the technical evaluation, special weightage was given to the functional requirements as well as to the technical support provided locally.

References

1. Michael Sherwood. Automation of the National Library. Paris : Unesco, 1984.
2. Kerry Webb. National Library of Malaysia: planning for inhouse computer facilities. Paris : Unesco, 1987.
3. Specification for Turnkey Integrated Library Computer System. Kuala Lumpur : National Library of Malaysia, 1987.

NATIONAL LIBRARY OF MALAYSIA
ORGANIZATION CHART


DIRECTOR GENERAL
DEPUTY DIRECTOR GENERAL
DIRECTOR OF SUPPORT SERVICES
DIRECTOR OF SERVICES
SUPPORT SERVICES
ACQUISITION OF NATIONAL COLLECTION
DOCUMENTATION
LIBRARY SERVICES
ADMINISTRATION
General Adm.
Personnel and Services
Finance
Building, Vehicles, Equipment Maintenance
Stock Control
Printing
TRAINING
Lib. Assist. Course I&II
Tech. Coop.Program
Seminars, Workshops, etc.
Attachments and Visits
CONSERVATION AND REPROGRAPHY
Conservation
Binding
Repairs
Reprography
PLANNING, RESEARCH, DEVELOPMENT
Advisory and Consultancy
Public Libraries
Special Libraries
Research
ACQUISITION
Malaysiana/ Reference Collection
Loan Collection
Serials Collection
LEGAL DEPOSIT GIFT, EXCHANGE
Legal Deposit Enforcement
Registration
Gift and Exchange
ISSN/ISBN
CATALOGUING
National Bibliography
Monographs
Serials
Audiovisuals
Cataloguing in Publication
INDEXING
Malaysian Periodicals
Malaysian Conferences
Malaysian Newspapers
Other Indexes (AGRIS)
Computerised Data Processing
REFERENCE
Information and Referral Services
Selective Dissemination of Information
Audiovisual Services
Collection Maintenance
CENTRE FOR MALAY MANUSCRIPTS
Identification and Acquisition
Documentation
Maintenance
LOANS
Membership Registration
Loan
Extension Activities
BRANCHES AND MOBILES
Branch I Sentul
Branch II Gombak
Branch III Bandar Tun Razak
Branch IV Labuan
Mobiles I, II, III



SYSTEM REQUIREMENTS

Integrated Library System
Acquisition
Cataloguing
Serial Management
OPAC
Circulation
Interlibrary loan
Conservation
ISBN
ISSN
SDI
Manuscript
Indexes

MINISIS Version H – Documentation

WandaJane Phillips
IDRC

MINISIS Background

MINISIS is a computerized textual database management system that can be used to create formats (database definitions) for the organization and maintenance of textual information. The strengths of MINISIS are in the organization of text and relationships between different texts rather than the manipulation of numbers for statistical purposes, although MINISIS is capable of some statistical operations.

MINISIS was originally developed to satisfy the needs identified by the Library at IDRC. The need for a similar system in developing countries was recognized and we acquired a mandate which surpassed our original intentions. Since then, MINISIS has developed through a great deal of interaction and association with its constituency (community of users).

The computer Systems Group, at IDRC, has examined the issues of MINISIS growth and development closely. Two items considered to be pressing issues were: the current software was not maKing use of the changes in HP architecture, and, the software development had not anticipated many of the enhancements made and future media for storage and retrieval. It was also becoming apparent that MINISIS could not fulfil its objectives unless it could migrate to other hardware platforms and operating systems.

ANALYSIS

The first stage of the documentation process is analysis. We are concerned, during this stage, with understanding the user needs and the proposed system.

Initially, we determined 'hat the user in developing countries would be our primary user. We have established a set of user profiles, or summaries from interviews with the MINISIS Outreach Officers and users in the Ottawa, Toronto, and Washington areas. We have extrapolated from these interviews to refine a questionnaire and develop an initial picture of the average user in our target audience.

With the identified audience, we plan to achieve two judgments: the types of tasks performed and the environment (physical job site and organizational methodology) in which they are performed. For the tasks, we looked at the functions of MINISIS, and asked the users to identify the functions and the series of functions most used. We also, through interviews, site visits, and the questionnaire, are looking to establish an idea about the milieu of MINISIS users. This information will tell us the strategies which may be effective for the documentation.

DESIGN

The issue in design is to develop a presentation which will effectively transmit the information to our audience. The development approach or method has been defined with the objective of satisfying particular goals; this is a flexible concept, as further exploration of our primary audience of users may reveal shortcomings in our assumptions. The physical design of the documentation is in process, although, like the development approach, certain assumptions and decisions have been made.

Development Verification

Following the design, writers will work with developers to identify the functional modules of information. These modules will then be written up, and will be put through four types of verification: accuracy, completeness, usability, and conformity. The accuracy of the documentation will be examined by software and specification developers and through prototyping. The completeness will be examined two ways: through inspection by developers and through prototyping. The prototyping will also reveal the usability of the document module. The final examination will scrutinize the module's compliance to the editorial standards and structural design for the documentation.

'two concepts should be explained to accompany my last point on verification: prototyping and usability. lb test modules we will use a tool which will act like MINISIS but will have nothing of the actual functionality of the software; this is a prototype of the interface. It is an empty shell or mock-up which can be made to act like MINISIS for the purposes of examining the interface design and the precision of the documentation. The second term, usability, refers to a measure of the success: can the user use the documents to answer questions, learn a task, or comprehend the system'?

The User

We are developing two task matrices: one generated from system point' of view, and one generated through responses to the questionnaire (specifically questions 1b and I.d). These matrices will help us identify the breakdown of tasks by user level (end user, database manager, programmer) and by result sought. The user level information will help us insure the appropriate information is in the appropriate document. The result sought or sequence of processes most commonly done (question I.d) will help us isolate the tasks for the users guides.

We want to understand the environment or milieu of our users so we can structure the physical, logical, and tactical aspects of the documentation to meet the needs of the users.

PHYSICALLY The current documentation may be too bulky and large for most installations.

LOGICALLY The divisions provided by IDRC do not reflect the partitioning preferred by many users; the functional units are not gathered in one place, nor are they clearly marked.

TACTICALLY The documentation does not answer questions clearly, nor does it identify all the steps and implications of actions in one place. Users do not always comprehend what the overall idea is and what the potential results are.

Profiles

END USER: The Primary end user of MINISIS is a person in a developing country who has little or no experience with computers or typewriters. The end user usually has completed a secondary level of education. The education level cannot be assessed, however, because of the diversity of education systems in developing countries. English is not the common first language of most end users, though most can speak English or

French as a second language. Generally, these users have to share office space and equipment. They may not have access to MINISIS documentation because the manuals are kept locked away or are simply not updated. Many users have received training and tend to use their classroom materials (e.g. hand-outs) as a reference to perform a task or solve a problem. If manuals are available, these users refer to Volume I. Their job titles are terminal operators or data entry clerks.

The Secondary end user is a person in a developed country who has some computer experience. These users have completed a secondary level of education or higher. Their first language is either English or French. Office space and equipment may have to be shared, but they always have either their own copy of the MINISIS documentation or easy access to one. Some of these users have received formal training from their MINISIS distributors, but the majority have not, generally because of budget and time constraints. These users are given a brief overview of MINISIS, instructions relating to their specific job functions, and a copy of the documentation and are expected to train themselves.

DATABASE MANAGER: The primary database manager is a person from a developing country who has some computer experience. They usually have attained a higher level of education than the primary end user and may have completed training on the use of the Hewlett-Packard (HP) 3300 series mini-computer. In many cases, though, this user has not received HP training because it is too expensive. A person with this l~ el of technical expertise in a developing country is in great demand, therefore, the rate of staff turnover for this position is high.

The secondary database manager is a person in a developed country who is computer literate. A large number of these users have educational backgrounds in Library Science. They may have completed HP training and be familiar with HP architecture.

PROGRAMMER: The primary programmer in a developing country usually has an educational background or experience in programming. Staff turn over is also a problem for this position.

The secondary programmer in a developed country has an educational background in Computer Science I Programming. These users are familiar with HP architecture through training or experience.

The Effect of Version H

Version H will be functionally compatible with Version G but the look and feel of the software is undergoing some radical changes. The function of the documentation in this situation is expanded from the usual role of reference material. Many of our current users will require retraining; we do not have the ability to accommodate all the retraining and training we expect Version H to require. This means the documentation must fill this gap; hence, we are developing a workbook which will be used by everyone, whether they are receiving training from IDRC or learning, or re.learning, on their own.

Since MINISIS Version H will come with more than one application and sites can modify or create an application for themselves, the documentation must be more generalist. The underlying principles must be made clear in order to enable users to operate a variety of applications while referencing one set of documents.

Many users want us to make the relationship between MINISIS and its operating platform (the HP 3000) clearer. Many users do not know when their results are affected by the HP system and when their results are affected by MINISIS. With the advent of the micro platform, and the possibility of future platforms, the documentation must look at ways of explaining where MINISIS ends and the operating system of the hardware platform begins. Where possible, we will identify the supported platforms and their contributions to the operations of MINISIS.

Our concerns when designing the documents are related to the analysis; what method or approach should we use, what logical divisions should we make, and what should the physical aspects of the documentation be?

The approach we use must satisfy the requirements already identified: ease of access, overall concepts, task descriptions, and concise directions. We expect the documentation to be modular, use conceptual models, be task-oriented, and minimalist in design.

MODULAR Modular documentation seeks to present information in

logical units. Each unit is concise and complete, representing a functional part

of the system.

CONCEPTUAL MODELS Conceptual modelling is a strategy for teaching a new idea or theory. The new information is presented using a familiar idea or process as a model for understanding the new idea or process.

TASK-ORIENTED In task-oriented instructions the entire sequence of events and requirements for a particular result are presented. The basic steps are listed, in order, with references to the other documents for more detailed information.

MINIMALIST Design

The documents will focus on real tasks and activities, stay away from lengthy, wordy explanations, and support error recognition and recovery. The structure of the information will encourage and guide exploration of the system and stay away from information overload.

The documentation, distributed on tape, has been helpful to some users, but we have found that many users do not print out the manuals from the tape. There are many reasons for this, but our primary concern is to find ways to give all users the same information in a timely way. The distribution of documentation is of greater importance for the release of Version H. The old documentation is not being revised for Version H, new documentation is being written, from the ground up.

This means the distribution methods must be examined, not only from the point of view of maintenance (updates), but also from the point of view of the initial release. We want to accommodate users who cannot afford to print out the documentation (either the initial release or the updates).

Usability Studies and Verification

We propose a series of usability studies be arranged and conducted during the course of the documentation cycle. This involves having working chunks of software, or simulations, and procedures documented. The test of the documents during this major reconstruction is crucial; we must determine if we have chosen a viable path for presenting the information to our audience. Of all the methods of usability studies available to us, we think a straightforward walk-through is the best bet.

The first item to understand is that any usability study involves people who match (to one degree or another) the primary and secondary audience. It is not feasible, given the constraints of our constituency and the time we have to develop and distribute MINISIS Version H, that we test the documents with actual users. We hope, instead, to match the abilities and backgrounds of the primary audience with local users and other easily obtained individuals (perhaps other employees at IDRC in Ottawa).

A walk-through consists of presenting a series of tasks and the appropriate documentation to a group of test subjects who resemble our primary and secondary audiences. The entire proceedings will be taped because we will ask the test subjects to speak aloud the mental processes they go through to complete the task. M' will also time each task and count the number of erroneous results. This will tell us whit mental models test subjects develop as they perform the task, how long the task can be expected to take (this can be compared to our concept of an acceptable limit), and the number of errors (which can also be compared to our concept of an acceptable limit). The test will also reveal areas in the documents where there is unclear or confusing text, and incomplete or inaccurate descriptions.

Another option, which could have more effective results, is called critical incident analysis. This involves using two terminals for a test; one terminal presents the software, the second terminal presents a questionnaire. Users are given a list of tasks to perform and documentation to help them; they are asked to report incidents, or events when they needed to refer to the documentation. In their report users identify whether the incident was a success or failure (did the documentation perform or not), what the details of the incident were, and how critical the incident was for the completion of the benchmark task. The tests can be repeated until a particular standard is achieved (a particular number or ratio of success and failure incidents, or the non-duplication of incidents in subsequent tests...).